AMERICAN

OURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

EDITED BY

E. W. CHITTENDEN
UNIVERSITY OF IOWA

A. B. COBLE
UNIVERSITY OF ILLINOIS

ABRAHAM COHEN
E JOHNS HOPKINS UNIVERSITY

G. C. EVANS
RICE INSTITUTE

F. D. MURNAGHAN
THE JOHNS HOPKINS UNIVERSITY

WITH THE COÖPERATION OF

FRANK MORLEY	HARRY BATEMAN	HARRY LEVY
E. T. BELL	J. R. KLINE	MARSTON MORSE
W. A. MANNING	E. P. LANE	ALONZO CHURCH

49

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

VOLUME LIV

1932

THE JOHNS HOPKINS PRESS
BALTIMORE, MARYLAND
U. S. A.

p16747

PRINTED BY THE J. H. FURST COMPANY
BALTIMORE, MD.

INDEX

PAGE

PAGE

ON THE CONSTRUCTION OF CYCLIC ALGEBRAS WITH A GIVEN EXPONENT.

CALCUTTA UNIVERSITY LIBRARY
Sec. No.............
Book No..........

By A. Adrian Albert.

1. *Introduction.* The most important linear associative algebras at present are the cyclic (Dickson) algebras. These algebras are normal simple algebras of order n^2 (degree n) over a non-modular field F. They are defined by a cyclic field Z of order n over F and a scalar γ in F. The exponent of an algebra A is the least integer ρ for which the direct power A^ρ is a total matric algebra, and it is desirable to know the conditions on γ that A have any given exponent. For algebras over an algebraic field Ω H. Hasse * has shown that the least integer σ for which γ^σ is the norm of a quantity of Z is the exponent of the cyclic algebra A. This result was also used to show that A is a division algebra if and only if

$$\gamma, \gamma^2, \cdots, \gamma^{n-1}$$

are all not the norms of any quantities of Z, a rather complicated set of conditions. For $n = 4$ the author recently simplified these conditions.†

The author has reduced essentially all problems on cyclic algebras to the case where $n = p^t$, p a prime.‡ He shows here that the exponent of a cyclic algebra A of degree p^t is an integer $\tau = p^e$ such that γ^τ is the norm of a quantity of Z while if $\pi = p^{e-1}$ then γ^π is not the norm of any quantity of Z. This reduces the number of powers to be considered from $n - 1$ to *two.* But a far greater simplification is obtained.

A cyclic field Z of order p^t over F contains a sequence of cyclic sub-fields

$$Z_t = Z, Z_{t-1}, \cdots, Z_1, Z_0 = F$$

where Z_i has order p^i over F. For each Z_i we may evidently define a norm symbol $N_i(\ )$. The author has now been able to secure the remarkable result that

$$\gamma^{p^e} = N(f), \qquad f \text{ in } Z,$$

if and only if γ itself has the property

$$\gamma = N_{t-e}(f_{t-e}), \qquad f_{t-e} \text{ in } Z_{t-e}.$$

* In his "Theory of Cyclic Algebras over an Algebraic Field," which will be published shortly in the *Transactions of the American Mathematical Society*.

† *Bulletin of the American Mathematical Society*, Vol. 37 (1931), pp. 301-312.

‡ See Theorem 5.

This implies that a cyclic algebra A of degree p^t has exponent p^e (the only possible type of value as is known) if and only if

$$\gamma = N_{t-e}(f_{t-e}), \qquad \gamma \neq N_{t+1-e}(f_{t+1-e}).$$

In particular a construction of all cyclic normal division algebras over an algebraic field Ω is furnished * when it is shown that algebra A of degree p^t over Ω is a division algebra if and only if γ is not the norm $N_1(f_1)$ of any quantity in the cyclic field Z_1 of order p over Ω.

2. *Elementary theorems.* We shall consider algebras over any non-modular field F. Let M be the algebra of all n-rowed square matrices with elements in F. If I is the n-rowed identity matrix and x is any matrix of M the equation

$$|\lambda I - x| = 0$$

is the characteristic equation $\phi(\lambda) = 0$ of x. Suppose that the equation $\phi(\lambda) = 0$ has no multiple roots. Then the invariant factors of x are $I_n \equiv \phi(\lambda), 1, \cdots, 1$ since each invariant factor I_j of x is divisible by I_{j-1} and $\phi(\lambda)$ is the product of the I_j. Hence *every* matrix y with the same characteristic equation as x has the same invariant factors as x and, as is well known,† is a transform zxz^{-1} of x by a non-singular matrix z of M.

Let y in M be commutative with x in M. If the characteristic equation of x has no multiple roots then y is a polynomial ‡ in x with coefficients in F. For if

$$\phi(\lambda) \equiv \lambda^n - \alpha_1\lambda^{n-1} - \cdots - \alpha_n = 0$$

is the characteristic equation of x then x is similar in F to

$$\begin{Vmatrix} 0 & 0 & \cdot & \cdot & \cdot & 0 & \alpha_n \\ 1 & 0 & \cdot & \cdot & \cdot & 0 & \alpha_{n-1} \\ 0 & 1 & \cdot & \cdot & \cdot & 0 & \alpha_{n-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & 1 & \alpha_1 \end{Vmatrix}$$

and, without loss of generality, may be taken to have this form. Let $\xi_1, \cdots, \xi_n$ be the scalar roots of $\phi(\lambda) = 0$, all distinct by hypothesis. The Vandermonde matrix

* This result was proved by the author for $p = t = 2$ (*Bulletin of the American Mathematical Society, loc. cit.*).

† Cf. L. E. Dickson's *Modern Algebraic Theories*, p. 104.

‡ A well known result in matrix theory. The author gives the proof here for its novelty and because it shows the actual form of the polynomial.

$$V = \| \xi_i^{j-1} \|, \qquad (i, j = 1, \cdots, n),$$

is then non-singular. By an elementary computation

$$Vx = \xi V, \qquad \xi = \left\| \begin{matrix} \xi_1 & & \\ & \ddots & \\ & & \xi_n \end{matrix} \right\|.$$

The matrix

$$t^{-1} = V'V = \| \sum_j \xi_j^{i+k-2} \| \qquad (i, k = 1, \cdots, n)$$

is non-singular and has elements in F. Let $\eta = VyV^{-1}$. Then

$$\eta = VyV^{-1} = VytV' = \| R(\xi_i, \xi_j) \| \qquad (i, j = 1, \cdots, n)$$

where if

$$yt = \| \beta_{ij} \|, \qquad (i, j = 1, \cdots, n),$$

then

$$R(\xi_i, \xi_j) = \sum_{s,t} \xi_i^{s-1} \beta_{st} \xi_j^{t-1} \qquad (i, j = 1, \cdots, n).$$

But $yx = xy$ if and only if $\xi\eta = \eta\xi$ so that

$$\| \xi_i R(\xi_i, \xi_j) - R(\xi_i, \xi_j) \xi_j \| = 0 \qquad (i, j = 1, \cdots, n).$$

Since $\xi_i = \xi_j$ only when $i = j$ we have $R(\xi_i, \xi_j) = 0$ $(i \neq j)$ and

$$\eta = \| R(\xi_i, \xi_i) \| = \left\| \begin{matrix} P(\xi_1) & & \\ & \ddots & \\ & & P(\xi_n) \end{matrix} \right\| = P(\xi)$$

where $P(\xi)$ has coefficients in F. But then $y = P(x)$.

Let A be a normal simple algebra of order n^2 (degree n) over F. Algebra M is a special case of A. As is well known $A = H \times D$ where H is a total matric algebra and D is a normal division algebra whose degree m is called the *index* of A. If ξ is any scalar root of the minimum equation of any quantity of grade m in D then, as was first proved by Wedderburn,*

$$A' = A \times F(\xi) = M \times F(\xi)$$

where M is a total matric algebra. Algebra A can then be thought of as an

* See the author's "On Direct Products," *Transactions of the American Mathematical Society*, Vol. 33 (1931), pp. 690-711, for reference and consequences of Wedderburn's result.

algebra over F whose quantities are n-rowed square matrices with elements in $F(\xi)$.

The characteristic equation of the general matrix of A is the so-called rank equation of A. If x in A has $\phi(\lambda)=0$ of degree n as its minimum equation with respect to F then obviously $\phi(\lambda)=0$ is the rank equation of A for x and is the characteristic equation of x. If moreover $\phi(\lambda)=0$ has no multiple roots then, as we have shown, the only matrices of A' commutative with x are polynomials in x with coefficients in $F(\xi)$. But then the only quantities of A commutative with x are polynomials in x with coefficients in F. Let next y in A have the same minimum equation as x. Then $y=z_0xz_0^{-1}$ where z_0 is in A', as we have seen. We may write

$$z_0=z_1+z_2\xi+\cdots+z_m\xi^{m-1}$$

where the z_i are in A. Since z_0x-yz_0 is zero and x and y are in the algebra A we have

$$\sum_i (z_ix-yz_i)\xi^{i-1}=0$$

whence $z_ix=yz_i$. It follows that

$$(\sum_i z_i\xi_i)x=y(\sum_i z_i\xi_i)$$

for any scalars ξ_i. The matrix $\sum_i z_i\xi_i$ is non-singular for the values $\xi_1=1$, $\xi_2=\xi,\cdots,\xi_m=\xi^{m-1}$ and hence the determinant of $\sum_i z_i\xi_i$ is not identically zero. It follows that there exist rational numbers $\eta_1,\cdots,\eta_m$ for which

$$z=z_1\eta_1+z_2\eta_2+\cdots+z_m\eta_m,$$

a quantity of A, is non-singular. Evidently $y=zxz^{-1}$, and we have proved

THEOREM 1. *Let A be a normal simple algebra of order n^2 (degree n) over F and let $\phi(\lambda)=0$ of degree n be the minimum equation of x in A. Suppose that $\phi(\lambda)=0$ has no multiple roots. Then the only quantities of A commutative with x are polynomials in x with coefficients in F. Every quantity y in A with the same minimum equation as x is a transform*

$$y=zxz^{-1}$$

of x by a non-singular quantity z of A.

Suppose that M is a total matric algebra of degree n over F. Algebra M has an *ordinary matric basis*

$$(e_{ij}) \qquad (i,j=1,\cdots,n),$$

with a multiplication table

$$e_{ij} \cdot e_{jk} = e_{ik}, \qquad e_{ij} \cdot e_{tk} = 0 \qquad (j \neq t;\ i, j, t, k = 1, \cdots, n).$$

Let y be in M and have

$$\psi(\lambda) \equiv \lambda^n - 1 = 0$$

as its minimum equation for F. The scalar roots of $\psi(\lambda) = 0$ are the ordinary n-th roots of unity and are all distinct. Consider also the quantity

$$y_0 = e_{12} + e_{23} + \cdots + e_{n1}$$

a matrix which is actually the so-called *rational canonical form* of y, and hence, by Theorem 1, similar to y. There thus exists a non-singular quantity z in M such that $y_0 = zyz^{-1}$. Let

$$\epsilon_{ij} = ze_{ij}z^{-1},$$

so that the (ϵ_{ij}) form a new ordinary matric basis of M. Then

$$y = z^{-1}y_0z = \epsilon_{12} + \epsilon_{23} + \cdots + \epsilon_{n1}.$$

THEOREM 2. *Let y be a quantity of a total matric algebra M of degree n over F, and let*

$$\psi(\lambda) \equiv \lambda^n - 1 = 0$$

be the minimum equation of y with respect to F. Then M has an ordinary matric basis

$$(e_{ij}) \qquad (i, j = 1, \cdots, n)$$

such that

$$y = e_{12} + e_{23} + \cdots + e_{n1}.$$

3. *Cyclic algebras.* Let $Z = F(i)$ be a cyclic field of order n over F and let $\theta(i)$ in Z be a root of the minimum equation $\phi(\lambda) = 0$ of i for F such that the n distinct roots of $\phi(\lambda) = 0$ in Z are $\theta^0(i) = \theta^n(i) = i$, $\theta(i)$, $\theta^2(i), \cdots, \theta^{n-1}(i)$. We may also represent the generator of the cyclic galois group of $\phi(i)$ by θ. We define an algebra A with the basis

$$(1) \qquad i^\alpha y^\beta \qquad (\alpha, \beta = 0, 1, \cdots, n-1),$$

and the multiplication table given by

$$(2) \qquad y^\beta i = \theta^\beta(i) y^\beta, \qquad y^n = \gamma \neq 0 \qquad (\beta = 0, 1, \cdots)$$

where γ is in F. Algebra A has been shown to be a normal simple algebra of degree n over F and is called a cyclic algebra.* We designate A by

* In the section on cyclic algebras of my paper "On Direct Products," *Ibid.*

$$(3) \qquad A = F(\gamma, Z, \theta).$$

Conversely every normal simple algebra A of degree n over F which contains a cyclic subfield Z of order n over F is a cyclic algebra, an immediate consequence of Theorem 1 for $y = \theta(i)$, $x = i$. The author has proved *

THEOREM 3. *The direct product*

$$(4) \qquad F(\gamma_1, Z, \theta) \times F(\gamma_2, Z, \theta) = M \times F(\gamma_1\gamma_2, Z, \theta),$$

where M is a total matric algebra.

Of particular importance in the theory of cyclic algebras is the structure of a cyclic algebra of degree p^t, p a prime, where the algebra is

$$A = F(\gamma^p, Z, \theta).$$

The cyclic field $Z = Z_t$ contains a cyclic sub-field $Z_{t-1} = F(u)$ of order $q = p^{t-1}$, the field of all quantities of Z symmetric in $\theta^q(i)$ and its $p-1$ iteratives. In fact Z is a cyclic field of order p over Z_{t-1}. Also

$$(5) \qquad yuy^{-1} = u\,[\theta(i)] = \theta_{t-1}(u), \qquad y^q u = uy^q.$$

Let

$$(6) \qquad z = \gamma^{-1}y^q$$

so that

$$(7) \qquad z^p = \gamma^{-p}y^n = \gamma^{-p}\gamma^p = 1.$$

Algebra A contains a cyclic sub-algebra B of order p^2 over $F(u)$ with the basis

$$(8) \qquad i^\alpha z^\beta \qquad (\alpha, \beta = 0, 1, \cdots, p-1)$$

and a multiplication table given by the minimum equation of i with respect to $F(u)$ and

$$(9) \qquad zi = \theta^q(i)z, \qquad z^p = 1.$$

But B is a normal simple algebra of prime degree p over $F(u)$ and is not a division algebra since $z^p = 1$ is impossible in a division algebra with a basis (8). Hence B is a total matric algebra. Moreover $\psi(\lambda) \equiv \lambda^n - 1$ is evidently the minimum equation of z with respect to $F(u)$ so that B has an ordinary matrix basis

$$(10) \qquad (e_{ij}) \qquad (i, j = 1, \cdots, n)$$

for B such that, by Theorem 2,

$$(11) \qquad z = e_{12} + e_{23} + \cdots + e_{n1}.$$

But then we may write

$$B = M_p \times F(u)$$

* In the section on cyclic algebras of my paper "On Direct Products," *Ibid.*

where M_p is a total matric algebra with a basis (10) over F and contains the quantity z of (11).

By the well known Wedderburn theorem

$$A = M_p \times D \tag{12}$$

where D is a normal simple algebra of degree $q = p^{t-1}$ over F. Algebra D is the algebra of all quantities of A commutative with all the quantities of M_p and contains the cyclic field Z_{t-1}. It follows that D is a cyclic algebra and has a basis

$$u^\alpha j^\beta \qquad (\alpha, \beta = 0, 1, \cdots, q-1), \tag{13}$$

where

$$j^\beta u = \theta^\beta{}_{t-1}(u) j^\beta, \qquad j^q = \delta \text{ in } F \qquad (\beta = 0, 1, \cdots). \tag{14}$$

But $yu = \theta_{t-1}(u)y$ so that $y^{-1}j$ is commutative with u.

The algebra of all quantities of A commutative with $F(u)$ is evidently B from the multiplication table (2) of A (with γ of course replaced by γ^p). Hence $y^{-1}j$ is in B. But since j is in D and z is in M_p the quantity j is commutative with z. Also y is commutative with $z = \gamma^{-1}y^q$. Hence $y^{-1}j$ in B is commutative with z. The only quantities of B commutative with z are polynomials in z with coefficients in $F(u)$ by (8) and (9). Hence $y^{-1}j = g(u, z)$, $j = yg = g[\theta_{t-1}(u), z]y = f(u, z)y$, where $f(u, z)$ is a polynomial in z with coefficients in F. We record this result in the form of a lemma.

LEMMA. *The quantity j of D has the form*

$$j = f(u, z)y \tag{15}$$

where $f(u, z)$ is a polynomial in z with coefficients in $F(u)$.

We now compute

$$\delta = j^q = N_{t-1}[f(u, z)]y^q \equiv f(u, z) \cdot f[\theta_{t-1}(u), z] \cdots f[\theta_{t-1}{}^{q-1}(u), z]y^q. \tag{16}$$

But $y^q = \gamma z$ so that (16) becomes

$$\delta\gamma^{-1} = N_{t-1}[f(u), z]z. \tag{17}$$

Let η be an indeterminate and let

$$g(\eta) = N_{t-1}[f(u, \eta)]\eta, \tag{18}$$

a polynomial in η with coefficients in F and finite degree. We can evidently write

$$g(\eta) = \sum_{i=0}^{r} g_i(\eta)\eta^{pi} \tag{19}$$

where the $g_i(\eta)$ are polynomials in η of degree at most $p-1$ with coefficients in F. Also we write

$$g_i(\eta)=\sum_{j=0}^{p-1} g_{ij}(\eta) \qquad (g_{ij} \text{ in } F). \tag{20}$$

Since $z^p=1$ we have

$$\gamma^{-1}\delta=g(z)=\sum_{i=0}^{r} g_i(z)=\sum_{j=0}^{p-1}\left(\sum_{i=0}^{r} g_{ij}\right)z^j. \tag{21}$$

But $1, z, \cdots, z^{p-1}$ are linearly independent with respect to F so that

$$\sum_{i=0}^{r} g_{i0}=\delta\gamma^{-1}, \quad \sum_{i=0}^{r} g_{ij}=0 \qquad (j=1,\cdots,p-1). \tag{22}$$

We now consider the polynomial

$$f(u)=f(u,1). \tag{23}$$

We have proved that

$$\begin{aligned} g(1)=N_{t-1}\left[f(u,1)\right]&=\sum_{i=0}^{r} g_i(1) \\ &=\sum_{j=0}^{p-1}\sum_{i=0}^{r}(g_{ij})=\sum_{i=0}^{r} g_{i0}=\delta\gamma^{-1}. \end{aligned} \tag{24}$$

The quantity

$$j_0=f(u)^{-1}j \tag{25}$$

differs from j by a polynomial in u with coefficients in F as a left factor and we may replace j in the basis of D by j_0. But

$$j_0{}^q=N_{t-1}\left[f(u)\right]^{-1}j^q=\gamma\delta^{-1}\delta=\gamma. \tag{26}$$

We have now given a short proof of the fundamental result of the paper. The further important theorems on cyclic algebras will be easily derivable from this proved result.

THEOREM 4. *Let $A=F(\gamma^p, Z, \theta)$ be a cyclic algebra of degree p^t over F, p a prime, so that Z contains a cyclic sub-field Z_{t-1} of order p^{t-1} over F with θ as generator of its cyclic galois group. Then*

$$A=M_p\times F(\gamma, Z_{t-1}, \theta_{t-1}) \tag{27}$$

where M_p is a total matric algebra of degree p over F.

4. *The norm condition.* The author has proved

THEOREM 5. *Let*

$$n=p_1{}^{b_1}p_2{}^{b_2}\cdots p_m{}^{b_m} \tag{28}$$

where the p_i are distinct primes. Every cyclic field Z of order n over F and with θ as a generator of its galois group is a direct product

$$Z = Z_1 \times Z_2 \times \cdots \times Z_m \tag{29}$$

of cyclic fields Z_i *of orders* $p_i^{b_i}$ *respectively and with* θ_i *respectively as generators, and conversely. Every cyclic algebra*

$$A = F(\gamma, Z, \theta) \tag{30}$$

of degree n *over* F *is a direct product*

$$A = F(\gamma_1, Z_1, \theta_1) \times F(\gamma_2, Z_2, \theta_2) \times \cdots \times F(\gamma_m, Z_m, \theta_m), \tag{31}$$

of cyclic algebras of degrees $p_i^{b_i}$ *respectively and with the same* $\gamma_i = \gamma$ *as* A. *Conversely every direct product* (31) *is a cyclic algebra* $F(\gamma, Z, \theta)$ *whose* γ *may be taken to be the* γ *of all the* $F(\gamma_i, Z_i, \theta_i)$. *Algebra* A *is a division algebra if and only if all the* $F(\gamma_i, Z_i, \theta_i)$ *are division algebras.*

The above theorem * reduces essentially all problems on cyclic algebras to the case where n is a power of a prime p.

We shall now give a short proof of the theorem that A $F(1, Z, \theta)$ is always a total matric algebra. By our proof of Theorem 4 this is certainly true when n is a prime p. Let the theorem be true for algebras of degree $r < n$. If n is not a power of a prime then Theorem 5 implies that A is a direct product of total matric algebras and hence is a total matric algebra. Let then $n = p^t$, p a prime. By Theorem 4 with $\gamma = 1$ we have $A = M_p \times F(1, Z_{t-1}, \theta_{t-1})$ is a direct product of total matric algebras and is a total matric algebra. The induction is complete and we have

THEOREM 6. *All cyclic algebras*

$$F(1, Z, \theta) \tag{32}$$

are total matric algebras.

We next give a short discussion of a necessary and sufficient condition that a general cyclic algebra be a total matric algebra. If γ is the norm $N(f)$ of a quantity f in Z then by replacing y in the basis (1) of A by $f^{-1}y$ we replace γ by $(f^{-1}y)^n = N(f^{-1})\gamma = 1$, so that $A = F(1, Z, \theta)$ is a total matric algebra. Conversely let $A = F(\gamma, Z, \theta)$ be a total matric algebra so that A is equivalent to $F(1, Z, \theta)$. Then A has a basis

$$i_0^{\alpha} y_0^{\beta} \qquad (\alpha, \beta = 0, 1, \cdots, n-1), \tag{33}$$

---

* For the results of this theorem see the author's "On Direct Products, Cyclic Division Algebras, and Pure Riemann Matrices," *Transactions of the American Mathematical Society*, Vol. 33 (1931), pp. 219-234.

such that

$$y_0 i_0 = \theta(i_0) y_0, \quad y_0^n = 1. \tag{34}$$

But i_0 has the same minimum equation as i, a cyclic irreducible equation. By Theorem 1 $i_0 = ziz^{-1}$ where z is in A. But then if

$$j = z^{-1} y_0 z \tag{35}$$

we have

$$j^n = 1, \quad ji = \theta(i) j. \tag{36}$$

Since $yi = \theta(i) y$ the quantity $j^{-1} y$ is commutative with i and is a polynomial in i with coefficients in F. But then $y = f(i) j$ and $y^n = N(f) = \gamma$ with f in Z.

THEOREM 7. *A cyclic algebra $F(\gamma, Z, \theta)$ is a total matric algebra if and only if γ is the norm $N(f)$ of a quantity f of Z.*

The exponent ρ of a normal simple algebra A over F is defined to be the least integer for which A^ρ is a total matric algebra. Let A be a cyclic algebra of degree n. As the author has shown

$$A^\rho = M^{\rho-1} \times F(\gamma^\rho, Z, \theta), \tag{37}$$

where M is a total matric algebra of degree n (*On direct products, loc. cit.*). But then $F(\gamma^\rho, Z, \theta)$ is a total matric algebra so that, by Theorem 7,

$$\gamma^\rho = N(f), \tag{38}$$

where f is in Z. Let σ be the least integer such that γ^σ is the norm of a quantity of Z. Then obviously $\sigma \leq \rho$. But $A^\sigma = M^{\sigma-1} \times F(\gamma^\sigma, Z, \theta)$ is a total matric algebra so that $\rho \leq \sigma$. Hence $\sigma = \rho$.

THEOREM 8. *The exponent ρ of a cyclic algebra is the least integer σ for which γ^σ is the norm of a quantity of Z.*

We now let $n = p^t$, p a prime. It is known that in this case ρ is a power of p (*On direct products, loc. cit.*), in fact a divisor of p^t. We thus have

THEOREM 9. *Let Z be a cyclic field of order p^t over F, p a prime, and let γ be in F. Then the least integer σ for which γ^σ is the norm of a quantity of Z is a divisor of p^t.*

Let $\tau = p^s$, $\pi = p^{s-1}$ and suppose that γ^τ is the norm of a quantity of Z while γ^π is not such a norm. Then $\tau \geq \sigma$, the least integer for which γ^σ is a norm. But if $\sigma = p^e$ then $e \leq s$. If $s = e + h$, $h > 0$ then $(\gamma^\sigma)^{p^{h-1}} = \gamma^{p^{e+h-1}} = \gamma^\pi$ is a norm, a contradiction. Hence $h = 0$ so that $\tau = \sigma$.

THEOREM 10. *If in Theorem 9 $\tau = p^s$, $\pi = p^{s-1}$ such that γ^τ is the norm of a quantity of Z while γ^π is not such a norm then $\sigma = \tau$.*

We shall now prove

THEOREM 11. *Let $Z = Z_t$ be a cyclic field of order p^t over F, p a prime, so that we can define a sequence of cyclic fields*

$$Z_t, Z_{t-1}, \cdots, Z_i, \cdots, Z_1, Z_0 = F \tag{39}$$

where Z_i is a cyclic field of order p^i over F with θ_i as a generator of its cyclic group. We may then also define a sequence of relative norm symbols $N_i(f_i)$ where f_i is in Z_i and

$$N_i(f_i) = f_i(u_i) \cdot f_i[\theta_i(u_i)] \cdots f_i[\theta_i^{p^i-1}(u_i)], \tag{40}$$

is in F. A quantity γ of F has the property that

$$\gamma^{p^e} = N_t(f_t) = N(f) \qquad f_t = f \text{ in } Z \tag{41}$$

if and only if the quantity γ itself has the property

$$\gamma = N_{t-e}(f_{t-e}) \tag{42}$$

for some f_{t-e} in Z_{t-e}.

We consider a cyclic algebra $A = F(\gamma, Z, \theta)$. Then if $\gamma^{p^e} = N(f)$ the algebra A^{p^e} is a total matric algebra. But, by Theorem 4,

$$A^p = H_1 \times F(\gamma, Z_{t-1}, \theta_{t-1}), \tag{43}$$

where H_1 is a total matric algebra. Similarly

$$A^{p^2} = H_1^p \times [F(\gamma, Z_{t-1}, \theta_{t-1})]^p = H_2 \times F(\gamma, Z_{t-2}, \theta_{t-2}), \tag{44}$$

and in general

$$A^{p^e} = H_e \times F(\gamma, Z_{t-e}, \theta_{t-e}), \tag{45}$$

where H_e is also a total matric algebra. But A^{p^e} is a total matric algebra so that $F(\gamma, Z_{t-e}, \theta_{t-e})$ is a total matric algebra. By Theorem 7 equation (42) is satisfied.

Conversely let (42) be satisfied. Then obviously from the definition of the field Z_{t-1} and similarly of the Z_i

$$\gamma^{p^e} = [N_{t-e}(f_{t-e})]^{p^e} = N_t(f_{t-e}),$$

so that our theorem is proved.

We shall now apply Theorems 9 and 10 to prove

THEOREM 12. *Let γ be the norm of a quantity in Z_{t-e} and not the*

norm of a quantity of Z_{t+1-e}. *Then the exponent of* $A = F(\gamma, Z, \theta)$ *of degree* p^t *over* F *is* p^e.

For if γ is the norm of f_{t-e} in Z_{t-e} then, as we have seen $\gamma^{p^e} = N_t(f_{t-e})$. But also $\gamma^{p^{e-1}}$ is not the norm of a quantity of Z since otherwise Theorem 11 would imply that $\gamma = N_{t+1-e}(f_{t+1-e})$, a contradiction of our hypothesis. Hence, if we put $\tau = p^e$, $\pi = p^{e-1}$ in Theorem 11, we obtain $\rho = \sigma = \tau$ as desired.

The consideration of γ itself instead of a power of γ is truly a remarkable simplification as working even with the case $p = t = 2$ has shown the author.

An immediate corollary of Theorem 12 and the fact that γ is in $Z_0 = F$ and hence the norm $N_0(f_0)$ of a quantity $f_0 = \gamma$ of F, is

THEOREM 13. *A cyclic algebra* $A = F(\gamma, Z, \theta)$ *of degree* p^t *over* F, p *a prime, has exponent equal to its degree if and only if* γ *is not the norm of any quantity of the cyclic field* Z_1 *of order* p *over* F, *a subfield of* Z, *that is*

$$\gamma \neq N_1(f_1), \quad f_1 \text{ in } Z_1.$$

Theorem 12 gave as simple a construction as is possible of algebras of degree p^t and exponent p^e, an *intermediate* power of p, while Theorem 13 gives a construction for the exponent p^t, the other end value being given by Theorem 7. In Theorem 7 the algebras were total matric algebras while in Theorem 13 they are actually division algebras. For Wedderburn showed * that A of degree n is a division algebra if γ^n is the least power of γ which is the norm $N(f)$ of a quantity f in Z, and hence, by Theorem 9, if n is the exponent of A. A simpler proof is perhaps the following. If A is a cyclic algebra $A = F(\gamma, Z, \theta)$ of degree and exponent n so that $A = H \times D$ where H is a total matric algebra of degree s and D is a normal division algebra of degree t then $st = n$. But $A^t = H^t \times D^t$ is a total matric algebra since H is a total matric algebra while, as is known (*On direct products, loc. cit.*) so also is D^t. Hence the exponent of A is at most t. But the exponent of A is n so that $t = n$ and A is a division algebra.

The above furnishes an essentially different proof of the Wedderburn norm condition. We have also proved.

THEOREM 14. *The algebras of Theorem 13 are cyclic normal division algebras over* F.

H. Hasse has considered cyclic algebras over algebraic number fields.

* Cf. L. E. Dickson's *Algebren und ihre Zahlentheorie*, p. 70, for Wedderburn's proof.

His algebraic theorems are evidently valid for any field $\Omega = R(\xi)$ where ξ is *any* quantity satisfying an equation with rational coefficients and irreducible in R, the field of all rational numbers. Hasse has proved that the exponent of any cyclic division algebra over a field Ω is equal to its degree. But now Theorems 13 and 14 give

THEOREM 15. *A cyclic algebra $A = (\gamma, Z, \theta)$ of order p^{2t} (degree p^t) over Ω, p a prime, is a division algebra if and only if*

$$\gamma \neq N_1(f_1)$$

for any f_1 in the cyclic sub-field Z_1 of order p over Ω, Z_1 contained in Z.

We have thus reduced the conditions for the construction of any cyclic normal division algebra over the most important type of algebra, to the condition *that a single quantity γ be not a norm of any quantity in a cyclic field of prime order over Ω.* Thus the general problem is of no more difficulty than the case where n is a prime after the author's reduction has been made. This indeed furnishes a construction of all cyclic normal division algebras over a field $R(\xi)$.

UNIVERSITY OF CHICAGO,
CHICAGO, ILL.

ON SELLING'S METHOD OF REDUCTION FOR POSITIVE TERNARY QUADRATIC FORMS.

By Burton W. Jones.

1. *Introduction.* Two quadratic forms are said to be *equivalent* when one may be transformed into the other by a linear transformation with integral coefficients and of determinant unity. A *reduced form* is a form representing a *class* of equivalent forms. It is defined by means of inequalities on the coefficients so that every form is equivalent to a reduced form and no two reduced forms are equivalent to each other.

One set of inequalities first found by Seeber and simplified by Eisenstein define what we shall in this paper refer to as an *Eisenstein-reduced* form. The proof that a form so defined has the properties mentioned above is geometric and, though simplified enormously and put on a firm foundation by L. E. Dickson * is still somewhat complex. However, the resulting reduced form is a very desirable one in that the leading coefficient a of the form is the least integer not zero represented by it and the second coefficient b is the least integer not ax^2 represented by the form, where x is an integer.

Selling † considers the form:

$$f = ax^2 + by^2 + cz^2 + 2gyz + 2hxz + 2kxy$$

and defines four new constants by the equations:

$$\begin{aligned} a + h + k + l &= 0 \qquad & b + g + k + m &= 0 \\ c + g + h + n &= 0 \qquad & d + l + m + n &= 0 \end{aligned}$$

and calls f reduced if g, h, k, l, m, n are all $\leqq 0$. Charve ‡ establishes this definition by proving that these conditions define an f uniquely except for permutations of g, h, k, l, m, n. Though Selling's proof followed geometric lines, Charve's is algebraic and this paper is based on the latter's work.

In view of the fact that permutations of g, h, k, l, m, n alter materially the form f, it is necessary to add further inequalities if one is to select

* L. E. Dickson, *Studies in the Theory of Numbers*, pp. 155-180. See this also for references on the Eisenstein inequalities which are given in paragraph 3 of this paper.

† E. Selling, "Des formes quadratiques binaires et ternaires," *Journal de Mathématiques* (3), Vol. 3 (1877).

‡ L. Charve, "Réduction des formes quadratiques ternaires positives," *Annales Scientifiques de L'École Normale Supérieure* (2), Vol. 9 (1880).

one form f from all those obtained by permutations of g, h, k, l, m, n. Then we will have a truly unique form. Borissow in his dissertation * considered this problem and constructed a table of reduced forms, according to Selling's definition, of determinant from 1 to 200. Though his supplementary inequalities were undoubtedly selected with a view to ease of computation, the resulting forms do not seem to be the best which can be obtained from Selling's inequalities. In view of the manifest advantages of the Eisenstein-reduced form it seems to the author highly desirable that the Selling reduced form should be made to correspond as closely with the Eisenstein reduced form as is consistent with Selling's original inequalities. That this is here accomplished is apparent from the discussion below.

The first part of this paper is devoted to the establishment, by algebraic methods based on those of Charve, of inequalities defining a unique reduced form and given in terms of the coefficients of f in paragraph 3 of this paper.

It should clearly be noted that, while Selling's and Charve's discussion yields a form † ϕ (see paragraph 2) unique except for permutations of the coefficients, the form f is *not* unique. For instance, the forms $x^2 + y^2 + z^2$ and $x^2 + y^2 + 3z^2 - 2xz - 2yz$ are both reduced under Selling's definition though they are equivalent. However, only the first of these forms satisfies inequalities (1), (2) and (3) of this paper. No two forms whose coefficients satisfy inequalities (1), (2) and (3) are equivalent unless they are identical, that is, unless corresponding coefficients are actually equal.

The third paragraph is devoted to a comparison of the above set and Eisenstein's set of defining inequalities. It is proved that every Eisenstein-reduced form for which g, h and k are not greater than zero is a Selling reduced form satisfying inequalities (1), (2) and (3) of paragraph 3—the closest correspondence possible, if Selling's original inequalities are to be retained. For the cases in which the forms do not coincide, a definite one-to-one correspondence is established between Eisenstein-reduced and Selling-reduced forms. This results in a stronger establishment of both theories, for the proof here contained may be thus considered as a proof of the Eisen-

* E. Borissow, "Reduction of Positive Ternary Quadratic Forms by Selling's Method, with a table of reduced forms for all determinants from 1 to 200," St. Petersburg 1890, 1-108 (Russian). The author here wishes to acknowledge his indebtedness to Professor J. V. Uspensky of Stanford University for the most generous loan of this book from his private library.

† Mordell, in *Proceedings of the Royal Society* (A), No. A 816, Vol. 131, p. 100, remarks that Selling found forms "leading to a *unique* reduced form." This is true only in the sense of the definition by Selling of a limited *class* of mutually equivalent forms any one of which may be considered to be the reduced form.

stein reduction inequalities or the proof of the Eisenstein reduction inequalities as a proof of the inequalities here.

In paragraph 4 the automorphs are listed.

2. *Supplementary inequalities.* We have associated with f a form

$$\phi = ax^2 + by^2 + cz^2 + dt^2 + 2gyz + 2hxz + 2kxy + 2lxt + 2myt + 2nzt$$
$$= -g(y-z)^2 - h(z-x)^2 - k(x-y)^2 - l(x-t)^2 - m(y-t)^2 - n(z-t)^2.$$

As Charve points out, ϕ represents the same integers as f.

Selling calls f reduced when g, h, k, l, m, n are all $\leqq 0$. Charve establishes this by first proving that a necessary and sufficient condition that none of g, h, k, l, m, n are positive is that

$$(a+b+c+d)/2 = -(g+h+k+l+m+n) = a+b+c+g+h+k$$

be a minimum under all equivalent transformations on f. Then, in order to show that ϕ is unique except for permutations of g, h, k, l, m, n he defines two associated transformations

$$s = \begin{pmatrix} \alpha & \alpha' & \alpha'' \\ \beta & \beta' & \beta'' \\ \gamma & \gamma' & \gamma'' \end{pmatrix} \quad \text{and } S = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \\ \beta_1 & \beta_2 & \beta_3 & \beta_4 \\ \gamma_1 & \gamma_2 & \gamma_3 & \gamma_4 \\ \delta_1 & \delta_2 & \delta_3 & \delta_4 \end{pmatrix}$$

where $\alpha''' = -\alpha - \alpha' - \alpha''$, $\beta''' = -\beta - \beta' - \beta''$, $\gamma''' = -\gamma - \gamma' - \gamma''$,

$\alpha = \alpha_1 - \delta_1$, $\alpha' = \alpha_2 - \delta_2$, $\alpha'' = \alpha_3 - \delta_3$, $\alpha''' = \alpha_4 - \delta_4$, $\beta = \beta_1 - \delta_1$, etc.

The application of S to ϕ is equivalent to the application of s to f and subtracting a number from all the elements of a column of S leaves s unaltered. We call S of "determinant 1" when s is of determinant 1.

For purposes of reference we state the following, which Charve proves,* as a lemma and reproduce a portion of his proof.

LEMMA. *Any transformation S, increases half the sum of the first four coefficients of ϕ by $-g(\sigma - 1)$ where*

$$2\sigma = (\beta_1 - \gamma_1)^2 + (\beta_2 - \gamma_2)^2 + (\beta_3 - \gamma_3)^2 + (\beta_4 - \gamma_4)^2$$

and correspondingly for h, k, l, m, and n.

If one denotes the coefficients of Φ, the transformed ϕ, by the corresponding capital letters we see $A = -g(\beta_1-\gamma_1)^2 - h(\gamma_1-\alpha_1)^2 - k(\alpha_1-\beta_1)^2 - l(\alpha_1 - \delta_1)^2 - m(\beta_1 - \delta_1)^2 - n(\gamma_1 - \delta_1)^2$ and similarly for B, C and D. Thus the coefficient of $-g$ in $A + B + C + D$ is 2σ.

* L. Charve, *loc. cit.*, p. 15.

COROLLARY.* *If* $g \neq 0$ *the only transformations leaving* $(A + B + C + D)/2$ *unaltered in value,* i. e. *leaving the form reduced by Selling's definition, are those in which* $\sigma = 1$ *above,* i. e. *those for which two of* $(\beta_1 - \gamma_1)$, $(\beta_2 - \gamma_2)$, $(\beta_3 - \gamma_3)$, $(\beta_4 - \gamma_4)$ *are zero and the other two* $+1$ *and* -1 *respectively; similarly for the other coefficients.*

The proof is given in Charve's paper.

We find it convenient to define coefficients a_{ij} as follows

$$
\begin{aligned}
&a_{11} = a, \quad a_{22} = b, \quad a_{33} = c, \quad a_{44} = d;\\
&b + 2g + c = a_{23} = a_{14} = a + 2l + d,\\
&a + 2k + b = a_{12} = a_{34} = c + 2n + d,\\
&a + 2h + c = a_{13} = a_{24} = b + 2m + d,\\
&2g = b_{23}, \quad 2h = b_{13}, \quad 2k = b_{12}, \quad 2l = b_{14}, \quad 2n = b_{34}, \quad 2m = b_{24}.
\end{aligned}
$$

Note $a_{ij} = a_{ji} = a_{kl}$ and $b_{ij} = b_{ji}$, where i, j, k, l is a permutation of 1, 2, 3, 4. Note also that $a_{11} + a_{22} + a_{33} + a_{44} = a_{12} + a_{23} + a_{13}$.

Our first condition on a reduced form then is Selling's

$$(1) \qquad b_{ij} \leqq 0, \quad \text{i. e.} \quad a_{ii} + a_{jj} \geqq a_{ij} \qquad (i, j = 1, 2, 3, 4).$$

To select a unique form out of each set of 24 forms found by permuting x, y, z, t we note that such a permutation merely permutes the subscripts of the a's (and b's) and the form is made unique by the following inequalities

$$(2.1) \qquad a_{ii} \leqq a_{jj} \quad \text{if} \quad i < j,$$

$$(2.2) \qquad \text{if } a_{ii} = a_{jj} \text{ with } i < j \text{ then } a_{jk} \geqq a_{ik} \quad \text{i. e.} \quad |b_{ik}| \geqq |b_{jk}|$$

where k is the lesser of 1, 2, 3, 4 not i nor j.

This is easily seen if we note that the interchange of i and j in $a_{jk} > a_{ik}$ reverses the inequality and if $a_{jk} = a_{ik}$, i. e. $a_{il} = a_{jl}$, such an interchange leaves the form unaltered, i. e. corresponds to an automorph.

Consider the transformation S_1 which is a particular form of S

$$S_1 = \begin{pmatrix} \alpha_1 - \delta_1 & \alpha_2 - \delta_2 & \alpha_3 - \delta_3 & \alpha_4 - \delta_4 \\ \beta_1 - \delta_1 & \beta_2 - \delta_2 & \beta_3 - \delta_3 & \beta_4 - \delta_4 \\ \gamma_1 - \delta_1 & \gamma_2 - \delta_2 & \gamma_3 - \delta_3 & \gamma_4 - \delta_4 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Note that no three-rowed determinant taken from the first three rows of S_1 is zero if s is non-singular for, from the definition of the α's, etc., $\alpha''' = l(\alpha, \alpha')$, $\beta''' = l(\beta, \beta')$, $\gamma''' = l(\gamma, \gamma')$ (where l denotes a linear function) implies $\alpha'' = l'(\alpha, \alpha')$, $\beta'' = l'(\beta, \beta')$, $\gamma'' = l'(\gamma, \gamma')$. Thus it follows that no column of S_1 is a linear combination of two others if s is non-singular.

* L. Charve, *loc. cit.*, p. 15.

Note also that S has the same effect on f and ϕ as the transformation S' obtained from S by multiplying each element of the matrix by -1. Thus if we are concerned merely with the resulting form we may use S or S' at pleasure, it being understood that if an explicit transformation s on f is desired we take that of S or S' for which s is of determinant 1. Hereafter we thus call S and S' the "same" transformation.

THEOREM 1. *If $ghklmn \neq 0$ the inequalities* (1) *and* (2) *define a unique form and the only automorphs are permutations of x, y, z and t.*

From the hypothesis and the corollary of the lemma it follows that every element of S_1 is 0, 1 or -1 and that every row as well as each set of differences between corresponding elements of two rows is $1, -1, 0, 0$ in some order. We thus have two possible types of S_1.

First, if three columns have three zeros and one column has one zero we see that each pair of columns having three zeros must have the element different from zero occurring in different rows. Consider two rows of the type

$$\begin{matrix} \epsilon & 0 & 0 & -\epsilon \\ 0 & \epsilon_1 & 0 & -\epsilon_1 \end{matrix}$$

the differences between corresponding elements being ϵ, $-\epsilon_1$, 0, $-\epsilon+\epsilon_1$ showing $\epsilon_1 = \epsilon$ and thus, after permutation of columns we have

$$S_1 = \begin{pmatrix} \epsilon & 0 & 0 & -\epsilon \\ 0 & \epsilon & 0 & -\epsilon \\ 0 & 0 & \epsilon & -\epsilon \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{where } \epsilon = 1 \text{ or } -1.$$

Second, if two columns have two zeros and two have three, we note that two non-zero elements occurring in the same column are equal as above and thus, after permuting columns and multiplying each element by -1 if necessary we have as the three rows $1, -1, 0, 0$; $0, -1, 0, 1$; $0, 0, -1, 1$ and the differences between the elements between the first and third are $1, 1, -1, -1$ which is not allowable.

Thus the only transformations permissible here are those obtained by permuting columns of S_1 displayed above, i. e.

$$S = \begin{pmatrix} \epsilon & 0 & 0 & 0 \\ 0 & \epsilon & 0 & 0 \\ 0 & 0 & \epsilon & 0 \\ 0 & 0 & 0 & \epsilon \end{pmatrix}$$

which is the identity transformation on ϕ or that obtained by multiplying

each element by -1. Thus the only permissible transformations merely permute x, y, z and t (or change the signs of all simultaneously). Such an interchange is barred by inequalities (2) unless it is an automorph and thus f is unique. We denote such a transformation by I.

THEOREM 2a. *If $2h = b_{13} = 0$ and $gklmn \neq 0$ the only transformations leaving $(a+b+c+d)/2$ unaltered are transformations I and those obtained by permuting the columns of*

$$T_{13} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The beginning sentence of the proof of theorem 1 holds here except that the difference between corresponding elements of the first and third rows need not be 0, 0, 1, -1 in some order. In fact, we will get new transformations only when it *is* not. However the sum of the differences must be zero. Under this hypothesis if we consider the last column to be that containing only one zero, $\alpha_4 - \delta_4 = \beta_4 - \delta_4 = \gamma_4 - \delta_4$ in the first division of the proof of theorem 1 and we see that we have no new transformation. The type excluded in the second division is precisely the type of T_{13} if we note that interchanging the first and third rows of T_{13} merely permutes the columns of T_{13} and multiplies each element by -1.

THEOREM 2. *If $b_{ij} = 0$ for only one pair of values i, j, the only transformations leaving $(a+b+c+d)/2$ unaltered are transformations I and those obtained by permuting the columns of T_{ij}, where T_{ij} is obtained from T_{13} by the following permutation of rows and columns*

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ i & k & j & l \end{pmatrix}$$

where $k < l$ and i, j, k, l are 1, 2, 3, 4 *in some order.*

This follows immediately from the above if we note that permuting the rows of any transformation S permutes the variables in the original form and the permutation of the columns of S permutes the variables in the resulting form.

Consider the following conditions on a form satisfying conditions (1) and (2)

$$\text{(3)} \qquad \begin{array}{l} \text{if } b_{ij} = 0 \text{ then } |b_{ik}| \leqq a_{ii} \text{ and } |b_{jk}| \leqq a_{jj}, \\ \text{that is, if } a_{ii} + a_{jj} = a_{ij}, \text{ then } a_{ik},\ a_{jk} \geqq a_{kk} \end{array}$$

where k is the lesser of 1, 2, 3, 4 not i nor j and l is the greater.

We prove

THEOREM 3. *When only one of g, h, k, l, m, n is zero, conditions* (1), (2) *and* (3) *determine a unique form to which every Selling reduced form is equivalent and T_{ij} (defined above) is an automorph if and only if one of the three equivalent equalities: $a_{ik} = a_{kk}$, $a_{ii} = -b_{ik}$, $a_{ii} = -b_{il}$ hold. Transformations I and T_{ij} are the only automorphs except possibly for certain permutations of the columns of T_{ij}.*

Under the transformation T_{13} we have the following correspondence

$$\begin{array}{llllllll} f & a_{11} & a_{22} & a_{33} & a_{44} & a_{13} = a_{24} & a_{23} = a_{14} & a_{12} = a_{34} \\ f' & a'_{11} & a'_{12} = a'_{34} & a'_{33} & a'_{23} = a'_{14} & a'_{13} = a'_{24} & a'_{44} & a'_{22} \end{array}$$

and thus under T_{ij} we have

$$\begin{array}{llllllll} f & a_{ii} & a_{kk} & a_{jj} & a_{ll} & a_{ij} = a_{kl} & a_{jk} = a_{il} & a_{ik} = a_{jl} \\ f' & a'_{ii} & a'_{ik} = a'_{jl} & a'_{jj} & a'_{kj} = a'_{il} & a'_{ij} = a'_{kl} & a'_{ll} & a'_{kk}. \end{array}$$

We consider f to satisfy conditions (1) and (2) and thus have $a'_{ik} \leqq a'_{jk}$.

If $a_{ik} < a_{kk}$ we have $a'_{kk} < a'_{ik}$ which, with $a'_{ll} = a'_{ik} + a'_{jk} - a'_{kk}$ implies $a'_{ll} > a'_{kj} \geqq a'_{ik} > a'_{kk}$ which shows (3) is satisfied by f', for variables corresponding to k and l are in their proper order.

If $a_{jk} < a_{kk}$ we have $a'_{ll} < a'_{ik} \leqq a'_{jk}$ which, as above, implies $a'_{jk} < a'_{kk}$ and thus f' with l and k interchanged satisfies (3), for $a_{ik} = a_{jl}$ and $a_{jk} = a_{il}$.

If $a_{ik} = a_{kk}$ then $a_{ll} = a_{jk} \geqq a_{kk}$, f satisfies (3) and T_{ij} is an automorph. Note also that $a_{ii} = -b_{il} = -b_{ik}$.

If $a_{jk} = a_{kk}$ then $a_{ll} = a_{ik} \geqq a_{kk}$, f satisfies (3) and T_{ji} is an automorph.

If, finally, a_{jk}, $a_{ik} > a_{kk}$ we see that a'_{ll}, $a'_{kk} > a'_{ik} = a'_{jl}$ and thus f' does not satisfy (3).

Thus, either f or f' but not both satisfies (1), (2) and (3) unless they are identical under a transformation T_{ij}. This completes the proof of theorem 3.

THEOREM 4a. *If $h = k = 0$, (i. e. $b_{13} = b_{12} = 0$) and $glmn \neq 0$ conditions* (1), (2), (3) *define a unique form f and the only new automorphs are K_{213} and possibly certain permutations of its columns where*

$$K_{213} = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = K_{312}.$$

We first justify the notation if we note that permuting the second and third rows and the second and third columns of the transformation leaves it unaltered.

Now consider transformations S_1 where $1, -1, 0, 0$ are, in some order, the elements of each of the rows and the set of differences between corresponding elements of the second and third rows. These conditions must not hold for the differences between the first and second or first and third rows if we are to have a transformation not already considered.

First, if three columns have three zeros and one has one zero permute columns if necessary to have only one zero in the last column. Note that $\mp 1, \pm 1, \pm 1$ are the first three elements of the last column and we have K_{213} above or the transformation obtained by multiplying each element by -1 or by permuting columns. Trial shows that K_{213} is an automorph.

Second, if two columns have two zeros and two have three, permute columns if necessary to make the first row $1, 0, 0, -1$. Since $b_{23} \neq 0$ and no column has three elements different from 0, two non-zero elements in the third or second column of the second two rows must be identical. Permute columns if necessary to put them in the third column and have

$$\begin{matrix} 1 & 0 & 0 & -1 \\ \cdot & \cdot & \epsilon & \cdot \\ \cdot & \cdot & \epsilon & \cdot \end{matrix}$$

Then either the first or last columns but not both have three zeros. Permute the first and fourth columns and multiply by -1 if necessary to make the first column contain three zeros. Permute the second and third rows if necessary to get

$$\begin{matrix} 1 & 0 & 0 & -1 \\ 0 & -\epsilon & \epsilon & 0 \\ 0 & 0 & \epsilon & -\epsilon. \end{matrix}$$

Since the difference between the first and third rows is not $1, -1, 0, 0$ in some order we see $\epsilon = -1$ and we have that all such transformations are obtained by permuting the columns of the following two transformations

$$D_{123} = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \text{ and } D_{132} = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & -1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Now, except for permutations of columns, $D_{123} = K_{213} \cdot T_{12}$ and $D_{132} = K_{213} \cdot T_{13}$ and, by the proof connected with conditions (3), D_{123} and D_{132} take f, satisfying (3), into a form which does not satisfy (3) unless T_{12} or T_{13} and therefore D_{123} or D_{132} are automorphs. This latter cannot happen because T_{12}, for example, is an automorph if and only if $-a_{11} = b_{13} = 0$ which is not true.

147

THEOREM 4. *If $b_{ij} = b_{ik} = 0$ and none of the other b's are 0, conditions* (1), (2), (3) *define a unique form f and the only new automorphs are certain permutations of the columns of K_{jik} where K_{jik} is obtained from K_{213} of theorem* 4a *by permuting the rows and columns according to the plan*

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ i & j & k & l \end{pmatrix}$$

and $K_{jik} = K_{kij}$.

In the first place $K_{jik} = K_{kij}$ since the permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ i & k & j & l \end{pmatrix} \text{ is } \begin{pmatrix} 1 & 3 & 2 & 4 \\ i & k & j & l \end{pmatrix} \text{ and } K_{312} = K_{213}.$$

Examining the changes made in any form f by K_{312} and noting the effect of the above permutation of 1, 2, 3, 4 we see that K_{jik} has the following effect on the coefficients of any form f: leaves b_{jk} unaltered, changes the sign of b_{ik} and b_{ij}, replaces b_{il} by $2b_{ik} + 2b_{ij} + b_{il}$, b_{jl} by $2b_{ij} + b_{jl}$, b_{kl} by $2b_{ik} + b_{kl}$ and increases the sum $(a + b + c + d)/2$ by $-(b_{ij} + b_{ik})$. K_{jik} is thus an automorph when $b_{ij} = b_{ik} = 0$.

Now the only transformations coming under the heading of this theorem are those obtained from K_{213}, D_{123} and D_{132} by the permutations indicated. Since T_{ij} increases the sum $(a + b + c + d)/2$ by $-b_{ij}/2$ for any form, it will increase the sum $(a + b + c + d)/2$ for the form resulting from the application of K_{jik} by $b_{ij}/2 \leqq 0$. Thus the total effect of $K_{kij} \cdot T_{ij}$ on the sum $(a + b + c + d)/2$ of a form is to increase it by $-b_{ik} - b_{ij}/2$. Since $K_{kij} \cdot T_{ij}$ leaves $(a + b + c + d)/2$ unaltered when and only when $b_{ik} = b_{ij} = 0$ it must be K_{jik}, D_{ijk} or D_{ikj} (except perhaps for permutation of columns) and since of these three only D_{ijk} increases $(a + b + c + d)/2$ by $-b_{ik} - b_{ij}/2$ (see the corollary to the lemma), we have the result that $K_{kij} \cdot T_{ij}$ must be obtained by a permutation of columns from D_{ijk}. Now, since K_{kij} is an automorph of f if $b_{ij} = 0 = b_{ik}$, the effect of D_{ijk} on such a form is the same as that of T_{ij}. But, from the proof of theorem 2, T_{ij} applied to f satisfying (1), (2) and (3) yields a form not satisfying (3), T_{ij} not being an automorph since $b_{ik} = 0 \neq a_{ii}$. Further, no permutation of variables in the resulting form can make it satisfy (1), (2) and (3). Thus D_{ijk} and, similarly, D_{ikj} can yield no new forms satisfying (1), (2) and (3); neither are they automorphs.

THEOREM 5a. *If $g = l = 0$* (i. e. $b_{23} = b_{14} = 0$) *and $hkmn \neq 0$, conditions* (1), (2) *and* (3) *define a unique form f and there are no new automorphs.*

Here in S_1 the second and third rows and the differences between the corresponding elements of the first and second rows and of the first and third are $1, -1, 0, 0$ in some order but this is not true of the first row or of the differences between corresponding elements of the second and third rows. The sum of the elements in any row is zero.

First, no element in the first row is ± 2. For, if it were, the elements in the same column of the rows below it would be ± 1 which implies that not more than one element of the first row is ± 2. If there were such a single element we could multiply every element of S_1 by -1 if necessary to make 2 an element of the first row. Note that the other elements in the first row would then be -1, -1 and 0 and permute columns if necessary to have

$$\begin{matrix} 2 & -1 & -1 & 0 \\ 1 & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot \end{matrix}$$

At least one of the elements in the last column is $\neq 0$ and therefore is -1. If -1 is the element in the last column of the second row the differences between corresponding elements of the first and second rows would be $1, -1, -1, 1$ which cannot be. The argument is the same for the third row.

Thus in the first row must occur two 1's and two -1's. Then, since $b_{12} \neq 0$ two of the elements in the second row must be directly below identical elements in the first. This holds for the third row also. Thus, after permutations of columns we have

$$\begin{pmatrix} 1 & -1 & 1 & -1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

which is singular. Thus results

THEOREM 5. *If $b_{ij} = b_{kl} = 0$ but none of the other b's are zero, conditions* (1), (2) *and* (3) *define a unique form f and there are no new automorphs.* (i, j, k, l *is a permutation of* 1, 2, 3, 4).

Note that $b_{ij} = b_{ik} = b_{jk} = 0$ implies no other b's are zero for $b_{ij} = b_{ik} = 0$ implies $a_{ii} = -b_{il} \neq 0$, etc.

THEOREM 6a. *If $g = h = k = 0$, conditions* (1), (2) *and* (3) *uniquely determine f and there are no new automorphs.*

Any of the first three rows of S_1 has $1, -1, 0, 0$ in some order, as its

elements but this is not true of the sets of differences between corresponding elements of any pair of rows from the first three. It can be seen from the proof of theorem 4a that the only case to consider here is that in which two columns have two zeros and two have three. Permute columns if necessary to make two non-zero elements occur in each of the first and third columns. Two elements in the same column must be of opposite sign and thus, by permuting the first three rows and multiplying each element by -1 if necessary we have for the first and third columns $1, -1, 0, 0$ and $0, 1, -1, 0$. Permute the second and fourth columns if necessary to make the element -1 of the first row occur in the second column, note that the last column cannot be all zeros and that

$$L_{123} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Thus, the only new transformations leaving $(a+b+c+d)/2$ unaltered are those obtained from L_{123} by permuting its columns and the first three rows. We note that interchanging the first and third rows of L_{123} merely permutes the columns and thus all transformations here to be considered are obtained by permutations of the columns of L_{123}, L_{132} and L_{213}.

Now, for any form f, L_{123} increases the sum $(a+b+c+d)/2$ by $-(b_{12}+b_{23}+b_{13}/2)$ and L_{ikj} increases it by $-(b_{ik}+b_{jk}+b_{ij}/2)$. It may be seen from the proof of theorem 4 that, for any form f, K_{ikj} increases $(a+b+c+d)/2$ by $-(b_{ik}+b_{jk})$ and leaves b_{ij} unaltered. Thus $K_{ikj}\cdot T_{ij}$ increases $(a+b+c+d)/2$ by $-(b_{ik}+b_{jk}+b_{ij}/2)$. Thus $K_{ikj}\cdot T_{ij}$ being a transformation leaving $(a+b+c+d)/2$ unaltered when and only when $b_{ik}=b_{jk}=b_{ij}=0$, must be L_{ikj} or obtained from it be a permutation of columns. We have three different transformations of this type for permutations of i, j, k. Now K_{ikj} is an automorph for f if $b_{ij}=b_{ik}=b_{jk}=0$ and thus L_{ikj} has the same effect as T_{ij}. Thus, if f satisfies conditions (3) as well as (1) and (2), its transformation by L_{ikj} does not satisfy (3). L_{ikj} cannot be an automorph because T_{ij} is an automorph only if $-b_{ii}=-b_{ik}=a_{ii}$ and $b_{ik}=0$ here.

This gives

Theorem 6. *If $b_{ij}=b_{ik}=b_{jk}=0$, conditions* (1), (2) *and* (3) *uniquely determine f and there are no new automorphs.*

Finally we prove

THEOREM 7. *Every positive ternary quadratic form is equivalent to a unique form f satisfying conditions* (1), (2) *and* (3).

This will follow from the theorems above if we prove

$$b_{ij} = b_{ik} = b_{jl} = 0 \text{ is inconsistent with } (1),\ (2),\ (3).$$

First note that it results from the hypothesis that

$$2a_{ii} = -b_{il},\ 2a_{jj} = -b_{jk} \text{ and thus } b_{il} \cdot b_{jk} \neq 0,$$
$$2a_{kk} = -b_{jk} - b_{kl},\ 2a_{ll} = -b_{il} - b_{kl} \text{ and thus } b_{kl} \neq 0$$

for, if it were, f would be expressible as $a_{ii}v^2 + a_{jj}u^2$ where v and u are each the difference of a pair of variables from x, y, z and t, which would mean that the determinant of f would be zero. Now, if f is to satisfy (1) and (2) we see that $i < l$ and $j < k$. Also $a_{kk} \leqq a_{ll}$ if and only if $-b_{jk} = 2a_{jj} \leqq -b_{il} = 2a_{ii}$. Noting that $b_{ij} = b_{ik} = b_{jl} = 0$ remains unaltered by the permutation $(ij)(kl)$ we see that the only values for i, j, k and l which need be considered are 1, 2, 4, 3 and 1, 3, 4, 2. In the first case $k = l = g = 0$, i. e. $a = -h$ which contradicts (3) with $i = 1$, $j = 2$. In the second case $h = l = g = 0$, i. e. $a = -k$ which contradicts (3) with $i = 1$, $j = 3$.

3. *A comparison of Selling- and Eisenstein-reduced forms*

$$f = ax^2 + by^2 + cz^2 + 2gyz + 2hxz + 2kxy.$$

We extend the Selling conditions for a reduced form henceforth in this paper by calling a form *Selling-reduced* if it satisfies not only conditions (1) but also (2) and (3). The form is unique from theorem 7. Translated into terms of the coefficients of f alone we have the following conditions for a Selling-reduced form f

(1.1) $\qquad g, h, k \leqq 0,$

(1.2) $\qquad a \geqq |h + k|;\quad b \geqq |g + k|;\quad c \geqq |g + h|$ *

(2.1) $\qquad a \leqq b \leqq c;\quad a + b + 2g + 2h + 2k - d - c \geqq 0.$

(2.2) $\qquad$ If $a = b$, $|g| \leqq |h|$; if $b = c$, $|h| \leqq |k|$;
if $a + b + 2g + 2h + 2k = 0$, $a \leqq |2h + k|$.

(3.1) $\qquad$ If $h = 0$, $a \geqq 2|k|$ and $c \geqq 2|g|$;
if $k = 0$, $a \geqq 2|h|$ and $b \geqq 2|g|$;
if $g = 0$, $b \geqq 2|k|$ and $c \geqq 2|h|$.

* The last condition is redundant in view of the inequalities below.

(3.21) If $a = -h - k$, then $a \geqq 2 \mid k \mid$ and $a + c + 2h \geqq b$,
i. e. $\mid h \mid \geqq \mid k \mid$ and $c + h \geqq b + k$.

(3.22) If $b = -g - k$, then $b \geqq 2 \mid k \mid$ and $b + c + 2g \geqq a$,
i. e. $\mid g \mid \geqq \mid k \mid$ and $c + g \geqq a + k$.

(3.23) If $c = -g - h$, then $c \geqq 2 \mid h \mid$ and $b + c + 2g \geqq a$,
i. e. $\mid g \mid \geqq \mid h \mid$ and $b + g \geqq a + h$.

We list the Eisenstein conditions for a reduced form *

(A) g, h, k all $\leqq 0$ or all > 0.

(B) $a \geqq 2 \mid h \mid, \quad a \geqq 2 \mid k \mid, \quad b \geqq 2 \mid g \mid.$

(C1) $a \leqq b \leqq c$, $[a + b + 2g + 2h + 2k \geqq 0]$.†

(C2) If $a = b$, $\mid g \mid \leqq \mid h \mid$; if $b = c$, $\mid h \mid \leqq \mid k \mid$;
[if $a + b + 2g + 2h + 2k = 0$, $a \leqq \mid 2h + k \mid$].†

(D1) For $g, h, k \leqq 0$: if $a = -2k$, $h = 0$;
if $a = -2h$, $k = 0$; if $b = -2g$, $k = 0$.

(D2) For $g, h, k > 0$: if $a = 2k$, $h \leqq 2g$;
if $a = 2h$, $k \leqq 2g$; if $b = 2g$, $k \leqq 2h$.

We prove the following lemmas relating the reduced forms of the two types.

LEMMA 1. *Every Eisenstein-reduced form with* $g, h, k \leqq 0$ *is Selling-reduced and only such are Selling-reduced.*

Consider such an Eisenstein-reduced form. Then, obviously all the conditions for a Selling-reduced form are satisfied if we note that (B) and (C1) imply $b \geqq 2 \mid k \mid, \quad c \geqq 2 \mid g \mid, \quad c \geqq 2 \mid h \mid$. No form with g, h and $k > 0$ can be a Selling-reduced form.

LEMMA 2. *A Selling-reduced form is Eisenstein-reduced if two or more of* g, h, k *are* 0 *or if none of the following mutually exclusive conditions are satisfied:*

I. $a < -2h$ *or* $a = -2h$ *and* $k \neq 0$.

II. $a < -2k$ or $a = -2k \neq -2h \neq 0$.

III. $b < -2g$ or $b = -2g$ *and* $k \neq 0$.

Inspection of the conditions shows that a Selling-reduced form is Eisen-

* L. E. Dickson, *loc. cit.*

† The condition in brackets is omitted if g, h, k are all > 0.

stein-reduced unless it fails to satisfy one of (B) or (D1). If $h = k = 0$ condition (3.1) shows that conditions (B) are satisfied and similarly for $g = h = 0$ and $g = k = 0$. I, II, III clearly include all cases in which one of (B) or (D1) is denied. To show that the three conditions are mutually exclusive note that I, (2.13) and (1.21) imply $b \geqq -2g - 2k \geqq -2g$, $b = -2g$ only if $k = 0$, $a \geqq -2k$ with the equality holding only if $a = -2h$. The proof is similar for II. Finally III and (2.13) implies $a \geqq -2h - 2k$, a not being equal to $-2k$ unless $h = 0$ nor to $-2h$ unless $k = 0$.

Lemma 3. *Let f' be a form with $a < -2h$ or $a = -2h$ and $k \neq 0$ and satisfying* (1) *and one of the following mutually exclusive sets of conditions:*

I′. $a \leqq b \leqq c$, *the first equality holding only if* $|g| \leqq |h|$, *the second equality not holding unless* $a = -2h = -2k$; *also* (2.13), (2.2), (3.1), (3.21).

II′. $a \leqq c \leqq b$, $h \neq k$, *the first equality holding only if* $|g| \leqq |k|$; $a + c + 2g + 2h + 2k \geqq 0$, *the equality holding only if* $a \leqq |2k + h|$; (3.1).

III′. $b < a \leqq c$, *the equality holding only if* $|g| \leqq |k|$; (2.13), (3.1) *and* (3.21) *the equality in* (2.13) *not holding.*

One and only one form f' may be obtained by a permutation of variables of a form f which is Selling-reduced but not Eisenstein-reduced. Conversely, one and only one form f, Selling-reduced but not Eisenstein-reduced, may be obtained by a permutation of variables of a form f'.

We first prove that for no form f' is $b = -g - k$ or $c = -g - h$ possible. Note that in each of I′, II′, III′ the sum of each pair of elements a, b, c is not less than $-(2g + 2h + 2k)$. This, with $a \leqq -2h$ implies $b \geqq -2g - 2k > -g - k$. Also $a + c + 2h \geqq -2g - 2k$ and, since in each of I′, II′, III′ is $c \geqq a$, $a + c + 2h \geqq 2a + 2h$. These, together, imply that $a + c + 2h \geqq a + h - g - k$, i.e. $c \geqq -h - g - k$. Thus $c > -g - h$ unless $k = 0$ which implies $a \geqq -2h$, $k = 0$, from (3.1), contrary to the conditions on f'.

If a form f occurs in I of lemma 2, its coefficients satisfy the inequalities of I'. $b \neq c$ from (2.22) unless $a = -2h = -2k$ since $a \leqq -2h$ and $a \geqq -h - k$ implies $|k| \leqq |h|$, the last equality holding only if the two previous ones hold. Thus, in view of the paragraph above, f in I and f' in I′ are identical. Furthermore, any permutation of variables alters $a \leqq -2h$

unless $a = -2h = -2k$ and the permutation is (yz). But (yz) alters $b \leqq c$ unless $b = c$ in which case it is an automorph.

If a form f occurs in category II permute y and z in f. This does not affect (1), converts II into I with $a \neq -2k$ if $a = -2h$, makes the changes noted in (2), the equality $c = b$ not considered since $|h| > |k|$ in any case, and (3.1) is left unaltered. Furthermore $a \neq -h - k$ from (3.21) since, before the permutation of y and z, $a < -2k$ or $a = -2k \neq -2h$. Interchange of y and z in conditions (3.22) and (3.23) interchanges the initial conditions $b = -g - k$, $c = -g - h$ and thus, from the first paragraph of this proof, none of (3.2) need be considered. Thus the conditions on f' in II′ are equivalent to the conditions II and (1), (2), (3) on f. Further permutation of the variables of f' alters the inequalities as above and our unique correspondence is established for this case.

If a form f occurs in category III note that (2.13) implies $a \geqq -2h -2k$, the equality holding only if $k \neq 0$ and therefore, from (2.23), $a + b + 2g + 2h + 2k \neq 0$. Also $a \neq b$ since $|g| > |h|$. Now permute x and y leaving (1) unaltered and making the indicated alterations in (2). (3.1) is left unaltered by the interchange and (3.22) becomes (3.21). As, in the previous paragraph, no other (3.2) need be taken into account. The correspondence is thus established as above.

I′, II′, III′ are categorical and mutually exclusive since I, II and III are.

LEMMA 4. *In every Eisenstein-reduced form*

$$F = Ax^2 + By^2 + Cz^2 + 2Gyz + 2Hxz + 2Kxy$$

with $G, H, K > 0$, the variables x, y, z may be permuted to give a unique form F' with $B - 2G \geqq A - 2H$, $H \geqq K \leqq G$ the first equality holding only if $A \leqq B$, the second only if $C \geqq B$, the third only if $C \geqq A$ and the last two only if $A \leqq B \leqq C$. Furthermore this form will satisfy conditions (B) *and* (D2) *together with those obtained by permuting x, y and z in the conditions. Every such form will fall into one of the three following mutually exclusive categories:*

I″. $A \leqq B \leqq A + C - 2H$, *the second equality not holding unless* $A = 2H = 2K$ *and* $B = C$.

II″. $A \leqq A + C - 2H \leqq B$, $A \neq H + K$.

III″. $B < A \leqq A + C - 2H$.

Furthermore no permutations of variables of the form F will yield two dis-

tinct forms falling into the same or different categories and one and only one form F may be obtained by a permutation of variables of a form F'.

Note that for every Eisenstein-reduced form with $H, K, G > 0$ not only do (B) and (D2) hold but also those conditions obtained from (B) and (D2) by permuting x, y, z with the corresponding change of designation of coefficients.

Consider a form F, Eisenstein-reduced with $G, H, K > 0$. Suppose one of G, H, K is less than the other two. Interchange variables if necessary until $H > K < G$. Then, without altering $H > K < G$, one may permute x and y if necessary to have a form for which $B - 2G > A - 2H$ or $B - 2G = A - 2H$ with $A \leqq B$.

If two of G, H, K are equal and less than the third we may permute variables to have a form for which $G = K < H$, the permutation being thereby uniquely defined except for an interchange of x and z.

a) If $B - 2G \geqq A - 2H$ and $C \geqq A$, we have it in the form desired.

b) If $B - 2G \geqq A - 2H$ and $C < A$, note $B - 2G = B - 2K > C - 2H$ and an interchange of x and z makes $B - 2G > A - 2H$ and $C > A$.

c) If $B - 2G < A - 2H$ but $B - 2G \geqq C - 2H$ note $C < A$, interchange x and z having $B - 2G \geqq A - 2H$ with $C > A$.

d) If, in any of the above cases, after the permutations noted we have $B - 2G = A - 2H$, $C \geqq A$ with $A > B$, interchange x and y to get $B - 2G = A - 2H$, $B > A$, $C \geqq B > A$ and $K = H < G$ in terms of the coefficients of the new form.

e) If $B - 2G < A - 2H$, $C - 2H$ interchange x and y and have $K = H < G$, $A - 2H < B - 2G$, $C - 2G$. These inequalities remain unaltered by a permutation of y and z and thus we can make $C \geqq B$.

If $G = H = K$, we may permute x, y and z so that $A \leqq B \leqq C$ and see that then $B - 2G \geqq A - 2H$. Thus we have shown that a permutation giving the required inequalities exists.

In order to show that this permutation is unique except for automorphs (i. e. the form is uniquely defined) suppose the coefficients of a form satisfy the requirements of lemma 4 (without I″ II″, III″). The interchange of x and y alters $B - 2G \geqq A - 2H$ unless $B - 2G = A - 2H$ and then alters $B \geqq A$ unless $A = B$ when $G = H$ and the form itself is unaltered. Similarly an interchange of x and z alters $K \leqq G$ unless $K = G$ and then $C \geqq A$ unless $C = A$; similarly for the interchange of y and z. The permutations

(xyz) and (xzy) leave $H \geqq K \leqq G$ unaltered if and only if $H = K = G$ in which case they alter $C \geqq B \geqq A$ unless $C = B = A$. If $A = B = C$ and $H = K = G$ both permutations are automorphs.

Finally, noting that, even after permutation of variables, $C \geqq 2H$ and therefore $A \leqq A + C - 2H$, we see that I″, II″, III″ form a set of categorical and mutually exclusive relationships between A, B and $A + C - 2H$. (In II″ $A \neq H + K$ for $A = H + K$ if and only if $A = 2H = 2K$, $C \geqq B$, $C \leqq B$, and therefore $C = B$, which is included in I″.)

THEOREM 8. *Every form f satisfying the conditions of lemma 3 is equivalent to a unique form F satisfying the conditions of lemma 4 and conversely. Furthermore, f occurs in category* I′ *if and only if F occurs in* I″ *and similarly for* II′ *and* II″ *as well as* III′ *and* III″.

First, suppose we have a form satisfying the conditions of lemma 3. Transform this form f by

$$\begin{pmatrix} -1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

and get $f' = ax^2 + by^2 + (a + c + 2h)z^2 - 2(g + k)yz + 2(a + h)xz - 2kxy$. Call this form F, denoting the coefficients by the corresponding capital letters. We note that if there is any correspondence by this transformation it must be I′ — I″, etc. It remains to prove that $f' = F$ satisfies the remaining conditions on F in lemma 4 together with (A), (D2), (B) and the conditions obtained from (B) and (D2) by permutations of the variables.

Conditions (A): $k \neq 0$ for $k = 0$ implies by (3.1) that $a \geqq -2h$. Thus $-g - k \neq 0$ and $a + h \neq 0$ since $a \geqq -h - k$.

Conditions (B) and those obtained from them by permuting x, y, z: $a \geqq 2(a + h)$ since $a \leqq -2h$; $a \geqq -2k$ similarly; $b \geqq -2g - 2k$ since $a \leqq -2h$ and since, for all of I′, II′, III′, (2.13) holds; therefore $b \geqq -2k$. Also $a + c + 2h \geqq -2g - 2k$ from (2.13); $a + c + 2h \geqq 2a + 2h$ since $c \geqq a$.

Conditions (D2) and those obtained from them by permuting x, y, z: If $a = -2k$ it is obvious that $a + h \leqq -2g - 2k$; similarly if $a = 2a + 2h$, $-k \leqq -2g - 2k$. If $b = -2g - 2k$, it follows from (1.2) that $-k \leqq 2a + 2h$. $b = -2k$ is impossible unless $g = 0$ for, from (2.1), $a \geqq -2h - 2g$; then $-g - k \leqq 2(a + h)$ from (1). If $a + c + 2h = 2a + 2h$, i. e. $c = a$, then $-g - k \leqq -2k$, i. e. $-g \leqq -k$ follows from the conditions for the equality $c = a$ in I′, II′ and III′. If $a + c + 2h = -2g - 2k$, it follows that $a + h \leqq -2k$ from the first equality and

(2.1) which imply that b is greater than or equal to the greater of a and c, which can happen only in II′ and when $a \leqq -2k-h$.

The conditions of lemma 4: (2.13) implies $B-2G \geqq A-2H$. Suppose the equality held with $A > B$, i. e. $a > b$. This latter occurs only in III′; but in that case the equality of (2.13) does not hold. In the second place, $K \leqq G$, i. e. $-k \leqq -g-k$ is true, the equality holding only if $g=0$ thus implying by (3.13) that $c \geqq -2h$ and therefore $a+c+2h = C \geqq A$. Thirdly, $K \leqq H$, i. e. $-k \leqq a+h$ by (1.21), the equality holding only if $a = -h-k$ and therefore when $C \geqq B$. Finally $K=H=G$ implies $g=0$, $a=-h-k$ and thus from (2.13), $b \geqq -h-k=a$ and thus $B \geqq A$ and $C \geqq B$ from the above.

Second, suppose we have the form F satisfying the conditions of lemma 4 together with (A), (B), (D2) and those obtained from (B) and (D2) by permutation of x, y, z. It is clear that, in virtue of (C), (B) and (D2) imply the conditions obtained from them by permutation of x, y, z and therefore, below, in citing these conditions we shall include the conditions so obtained from them.

Transform F by

$$\begin{pmatrix} -1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

(the inverse of the previous transformation) and get

$$F' = Ax^2 + By^2 + (A+C-2H)z^2 - 2(G-K)yz - 2(A-H)xz - 2Kxy.$$

We prove $F'=f$ satisfies the conditions I′, II′ or III′ and $a < -2h$ or $a = -2h$ with $k \neq 0$ and that f satisfies (1).

$a \leqq -2h$ the equality holding only when $k \neq 0$ if and only if $A \leqq 2A - 2H$, i. e. $2H \leqq A$, $K \neq 0$ which follows from the conditions given.

(1.1) hold, for $G-K \geqq 0$, $A-H > 0$, $K > 0$.

(1.2): $A \geqq A-H+K$, i. e. $H \geqq K$; $B > G-K+K = G$; $A+C-2H > G-K+A-H$, i. e. $C-H > G-K$.

(2.13): for all permutations it is true that $A+B$, $2A+C-2H$, $B+A+C-2H \geqq 2G+2A-2H$. This is true since $B-2G \geqq A-2H$. The equality can hold only if $B-2G = A-2H$ or II″ and $C=2G$.

(2.21) and (2.22) for I′: if $A=B$, $G-K \leqq A-H$, i. e. $A \geqq G+H-K$ since $H \geqq G$ and $A \geqq 2H$; if $B = A+C-2H$, $A-H \leqq K$ for the first equality by I″ implies $A=2H=2K$.

(2.21) and (2.22) for II′: if $A=A+C-2H$, $G-K \leqq K$ is true since the first implies $C=2H$ and (D2) applies. $K \neq A-H$ is true.

(2.22) for III′: $A = A + C - 2H$, $G - K \leqq K$ as above.

(2.23): for I′ we have the equality only if $B - 2G = A - 2H$; then $A \leqq 2A - 2H + K$ since $A \geqq 2H - K$. For II′ we have the equality only if $C = 2G$; then $A \leqq 2K + A - H$, i. e. $H \leqq 2K$ from (D2).

(3.1): The only possible equality in (1.1) above is $G = K$, i. e. $g = 0$. Then $B \geqq 2K$ and $A + C - 2H \geqq 2A - 2H$, i. e. $C \geqq A$, the latter being true in virtue of lemma 4 with $G = K$.

(3.21): For I′ the only equality of (1.2) is $H = K$ above. $A \geqq 2K$ and $a + c + 2h = C \geqq B$ is true from the above. For III′ with $H = K$ we have the same.

To complete the proof call T the transformation of the second part of the proof of this theorem. T takes a form F satisfying the conditions of lemma 4 into a form f satisfying the conditions of lemma 3. Suppose that there is another equivalent transformation T' (that is, with integral coefficients and of determinant unity) taking F into $f' \neq f$ satisfying the conditions of lemma 3. Then T^{-1} would take f' into $F' \neq F$ where F' satisfies the conditions of lemma 4. Then $f' \sim f$ and $F' \sim F$. By lemma 4, F' and F are equivalent to two forms F'_1 and F_1, $F'_1 \neq F_1$, which are Eisenstein-reduced. By lemma 3, f' and f are equivalent to two forms f'_1 and f_1, $f'_1 \neq f_1$, which are Selling-reduced. This occurrence contradicts the proof that no two Eisenstein-reduced forms are equivalent as well as the corresponding proof for Selling-reduced forms.

Thus we have also proved the

COROLLARY. *No two Selling-reduced forms are equivalent if and only if no two Eisenstein-reduced forms are equivalent.*

Thus the two theories of reduction are correlated.

It is interesting to note here that Charve proves (*op. cit.*, p. 18) that the three least integers* not necessarily distinct represented by a Selling-reduced form are included among the following:

$$a,\ b,\ c,\ a + b + c + 2g + 2h + 2k,\ a + b + 2k,\ a + c + 2h,\ b + c + 2g.$$

Since $a \leqq b \leqq c \leqq a + b + c + 2g + 2h + 2k$ we may exclude the fourth quantity in the list.

4. *Automorphs.* The automorphs of forms satisfying Eisenstein's inequalities are found in *Studies in the Theory of Numbers* (*op. cit.*). We accordingly find here merely the automorphs for those forms whose coefficients satisfy inequalities (1), (2) and (3) but not all of (B) and (D1), i. e. which occur in one of the categories I, II, III of the last section.

* Exclusive of x^2a, $ax^2 + by^2 + 2bxy$, etc.

The ten cases on page 180 of *Studies* in which none of g, h, k (that is, r, s, t) are zero correspond to the automorphs which are permutations of x, y, z and the fourth variable t. It may be quickly verified that of these only the following can occur for a Selling reduced form [that is, satisfying (1), (2) and (3)] which is not Eisenstein-reduced:

	Conditions on the form	Automorphs
1_1	$a+2h+k=b+2g+k=0$	$(z-x,\ z-y,\ z)$
1_2	$a=b,\ g=h$	$(-y,\ -x,\ -z)$
1_3	$b=c,\ h=k$	$(-x,\ -z,\ -y)$
1_4	$a=b,\ a+b+2g+2h+2k=N=0$	$(y-z,\ x-z,\ -z)$
1_5	$a=b=c,\ N=0$	$(z-y,\ -y,\ x-y)$
		$(-x,\ z-x,\ y-x)$
1_6	$a=b=c,\ N=0,\ g=h$	$(x-z,\ x,\ x-y)$
		$(y,\ y-z,\ y-x)$
1_7*	$a=b=c,\ N=0,\ h=k$	$(y-z,\ y-x,\ y)$
		$(x,\ x-y,\ x-z)$
		$(z-y,\ z,\ z-x)$.

If one and only one of g, h, k, l, m, n is zero we prove that we have to consider merely the following five conditions on the form and the corresponding automorphs:

	Conditions on the form	Automorphs
2_1	$b=-2k,\ g=0\neq h$	$(-x,\ y-x,\ -z)$
2_2	$c=-2h,\ g=0\neq k$	$(-x,\ -y,\ z-x)$
2_3	$c=-2g,\ h=0\neq k$ (thus $a\neq b$)	$(-x,\ -y,\ z-y)$
2_4	$a=-2k=-2h$	$(x-y-z,\ -y,\ -z)$
2_5	$a=-h-k,\ b+k=c+h$	$(y+z-x,\ z,\ y)$.

Proof. Consider the cases of theorem 3. The three cases $a=-2k$, $h=0$; $a=-2h$, $k=0$; $b=-2g$, $k=0$ occur only when the form is Eisenstein-reduced since $a=-2k$ implies $b\geqq -2g$ and $b=-2g$ implies from (2. 2) that $a\leqq -k$ which is false, etc. $b=-2k=-2g$ implies, from (2. 1) that $a\geqq -2g-2h>b$ unless $h=0$ and $a=-2g=b$ when, from (2. 2), $|g|\leqq 0$ which is false. The same argument holds for $c=-2h=-2g$. Also $b=-g-k$ implies $a\geqq -g-k-2h>b$ unless $h=0$; and a similar argument holds for $c=-g-h$. The form of the automorph for each remaining case can be obtained from theorem 3.

We have already shown (lemma 2) that a Selling-reduced form is

* This can occur only in combination (iv) below.

Eisenstein reduced if two of g, h, k are zero. Suppose $b_{ij} = b_{ik} = 0$ where one of i, j, k is 4 and none of the other b's, except perhaps b_{jk}, is zero. Then $2a_{ii} = -b_{il}$, $2a_{ll} = -b_{il} - b_{kl} - b_{jl} > 2a_{ii}$. Thus $l > i \neq 4$. Thus we have to consider only $b_{ij} = b_{i4} = 0$. Note that, from theorem 3, $|b_{ik}| \leqq a_{ii}$ and $|b_{jk}| \leqq a_{jj}$ implies $|b_{il}| \geqq a_{ii}$ and $|b_{jl}| \geqq a_{jj}$ when $b_{ij} = 0$. Then $|b_{i4}| \geqq a_{ii}$ where $b_{i4} = 0$, which is impossible.

Thus, except for the identity, the only automorphs for any form are those above and, in certain cases, those obtained by certain permutations of the columns of automorphs 2_1 to 2_5. Investigation reveals that for any form Selling-reduced but not Eisenstein-reduced, one of three things may occur: first, none of the conditions 2_1 to 2_5 hold in which case the automorphs are given by 1_1 to 1_6; second, none of the conditions 1_1 to 1_6 hold and not more than one of 2_1 to 2_5 hold, in which case the automorphs are as given; third, one of the combinations below holds.

	Conditions on the form	Automorphs
(i)	$a = -2h = -2k,\ b = c$	1_3, 2_4, 2_5.
		$(2_4 \cdot 1_3 = 2_5$ and $2_5 \cdot 1_3 = 2_4$
(ii)	$a = b = -h - k,\ c = -2h,\ g = 0$	1_4, 2_2, 2_5, $2_2 \cdot 1_4$, $2_5 \cdot 1_4$.
(iii)	$a = b = -2k = -2h,\ g = 0$	1_4, 2_1, 2_4, $2_1 \cdot 1_4$, $2_4 \cdot 1_4$.
(iv)	$a = b = c = -2k = -2h = 2,\ g = 0$	See below.

In the last case above which includes only one primitive form transform the form by the first transformation of lemma 5 to get an Eisenstein-reduced form. Its automorphs are given in *Studies*. Note that no pair of (i), (ii), (iii) can hold without (iv) holding.

We give the following multiplication table for the combinations above:

$A =$	2_1	2_2	2_4	2_5
$A \cdot 1_4 =$	$(z-y, x-y, z)$	$(z-y, z-x, -y)$	$(y+z-x, z-x, z)$	$(x-y-z, -z,$

From the above information the automorphs of any form may be found if it is Selling-reduced but not Eisenstein-reduced.

A NEW TYPE OF ARITHMETICAL INVARIANCE.

By E. T. BELL.

1. The arithmetical invariance in question will be more easily seen from an example than from formal definitions. The following notation is fixed.

$m \equiv$ an arbitrary constant odd integer > 0.

μ, ν, δ, τ, with suffixes, are variable integers; the μ, δ, τ are odd, the ν odd or even; the δ, τ are > 0, the $\mu \gtrless 0$, and the $\nu \gtreqless 0$.

$$\Phi_{si} \equiv \sum_{a=1}^{2(m^2-1)-4s+4i} \mu^2_{sia} + 4 \sum_{b=1}^{2(m^2-1)-4i} \nu^2_{sib} + 2 \sum_{c=1}^{2s+1} \delta_{sic}\tau_{sic},$$

$$(s = 0, \cdots, (m^2-1)/2;\ i = 0, \cdots, s).$$

A sum $\sum_p^q$ in which $q < p$ is vacuous and is to be ignored.

$g(z)$ is uniform and finite for all integer values of z; $g(z) = -g(-z)$; $g(0) = 0$; $g(z)$ is otherwise arbitrary.

Each of $e_1, \cdots, e_{2s+1}$ is a definite one of $+1$ or -1, and $\sum_e$ refers to the 2^{2s+1} values of the matrix $(e_1, \cdots, e_{2s+1})$.

N is an arbitrary integer $\geqq 0$.

$A_{si}(m)$, $B_{si}(m)$ are integers, and depend upon m alone.

The invariance considered refers to the $A_{si}(m)$, $B_{si}(m)$ (in this example), and concerns functions $g(z)$ whose arguments z are certain linear, homogeneous functions of all the τ's obtained from the set of all representations of N simultaneously in the system of $(m^2+1)(m^2+3)/8$ forms Φ_{si} defined above. Each form Φ_{si} is in $2(2m^2-1)$ indeterminates μ, ν, δ, τ. The total number of $A_{si}(m)$, $B_{si}(m)$ in this *invariant set* for the system of forms Φ_{si} is $(m^2+1)(m^2+3)/4$. The invariance is of the kind that the linear relation between the functions $g(z)$ just described has as its coefficients the integers in the invariant set, *and this relation is the same for all N represented in the system of forms Φ_{si}*. The trivial invariant set in which each $A_{si}(m)$, $B_{si}(m)$ is zero is excluded; *there is precisely one non-trivial invariant set.* The same invariant set appertains to an infinity of distinct systems of forms. We proceed to state the linear relation and to indicate how the corresponding invariant set may be found.

With respect to the form Φ_{si}, for fixed (s, i), we construct first the sums

$$g_{sim}(N) \equiv \sum g\left(\sum_{\sigma=1}^{2s} e_\sigma \tau_{si\sigma} + m e_{2s+1} \tau_{si\,2s+1}\right),$$

$$g_{si1}(N) \equiv \sum g\left(\sum_{\sigma=1}^{2s} e_\sigma \tau_{si\sigma} + e_{2s+1} \tau_{si\,2s+1}\right),$$

where $\sum$ refers to all solutions of $\Phi_{si} = N$, and $(e_1, \cdots, e_{2s+1})$ $(s \geqq 0)$ is the general matrix of e's. Write

$$g_{si}^{(m)}(N) \equiv \sum_e e_1 \cdots e_{2s+1}[A_{si}(m) g_{sim}(N) + B_{si}(m) g_{si1}(N)],$$

where $\sum_e$ (as stated above) refers to the 2^{2s+1} e-matrices, and where the A, B are for the moment arbitrary constants. Thus $g_{si}^{(m)}(N)$ is defined for Φ_{si}, (s, i) fixed. Finally we sum $g_{si}^{(m)}(N)$ over all forms Φ_{si} in the system of $(m^2+1)(m^2+3)/8$.

THEOREM. *There is precisely one non-trivial* ($\equiv$ *not all zero*) *invariant set of* $(m^2+1)(m^2+3)/4$ *integers*

$$A_{si}(m), B_{si}(m) \qquad (s = 0, \cdots, (m^2-1)/2;\ i = 0, \cdots, s)$$

such that

$$\sum_{s=0}^{(m^2-1)/2} \sum_{i=0}^{s} g_{si}^{(m)}(N)$$

vanishes for all integers $N \geqq 0$; *the integers in the invariant set do not depend upon* N.

As already remarked, the *invariance* consists in the fact that the unique invariant set is the same for all $N \geqq 0$. The following features of such theorems are noteworthy: the *uniqueness* of the invariant set; the fact that the number of indeterminates in each form of the system concerned is of the *second degree* in the given constant integer (here m), namely $4m^2 - 2$ here; the arbitrariness of the function g; the existence of *an infinity* of systems of forms having the same invariant set.

The above set is based on an arbitrary *odd* positive integer m, and an arbitrary *odd* function $g(z)$, for the particular system of forms Φ_{si}. Invariant sets also exist for even m, or for unrestricted m, and for even functions g or for completely arbitrary functions.

The distinctive feature of such invariant sets is the *form of the number of indeterminates*, namely, *a polynomial of degree* 2 in the *given integer* (m above).

2. The origin of the invariant set in § 1 can be briefly indicated. Write

$$\lambda(x) \equiv \Phi_2\vartheta_3\vartheta_1(x)/\vartheta_0(x), = 4\sum q^{t/2}\left[\sum \sin \delta x\right],$$

where the theta notation is that of Jacobi, and the first $\sum$ refers to all odd integers $t > 0$, the second to all positive divisors, δ, of t. Then, for m an odd integer > 0, as in § 1, we have

$$\lambda(mx)/\lambda(x) = G_1^{(m)}(w)/G_0^{(m)}(w),$$

where $w \equiv \lambda^2(x)$, and the G's are polynomials in w, k^2, where k is the modulus of $\lambda(x)$,

$$G_1^{(m)}(z) = m + \sum_{s=1}^{(m^2-1)/2} a_{1s}^{(m)} w^s,$$

$$G_0^{(m)}(z) = 1 + \sum_{s=1}^{(m^2-1)/2} a_{0s}^{(m)} w^s,$$

in which $a_{js}^{(m)}$ $(j = 0, 1)$ is a polynomial in k^2 of degree at highest $t - 1$ in k^2, with rational integer coefficients. Assuming these polynomials known, we use $k^2 = \vartheta_2^4/\vartheta_3^4$, and proceed to paraphrase the resulting identity.*

As similar reductions have been explained in detail in the places cited, and elsewhere, we shall omit them, especially as they add nothing to the main object of this note, which is to call attention to a remarkable new type of arithmetical invariance.

Considerably more may be said about the integers in the invariant set of the Theorem than is there stated, and similarly for all such sets. In a paper to be published elsewhere I have shown how these integers can be given explicitly as polynomials in m.

The uniqueness of the invariant set is an immediate consequence of the like (easily proved by a simple contradiction, if not already classic) for the formulas giving the real multiplication of elliptic functions.

The infinity of systems of forms having the same invariant set follows by paraphrasing any of the identities, equivalent to that obtained from $\lambda(mx)/\lambda(x)$, by introducing any common factor, composed of thetas or

---

* See, for example, *Transactions of the American Mathematical Society*, Vol. 22 (1921), p. 1; or "Algebraic Arithmetic," *Colloquium Publications of the American Mathematical Society*, Vol. 7 (1927). In the last refer particularly to p. 48, (10. 4), for the appropriate e-theorem concerning the separation of the products of sines, arising from the powers of $\lambda(x)$, into the corresponding sums. From these we write down the integers A, B of the Theorem in § 1, as $\pm$ the respective integer coefficients in the polynomials $a_{js}^{(m)}$.

elliptic functions, into the numerator and denominator of the expression for $\lambda(mx)/\lambda(x)$. At any stage the elliptic functions can be replaced by their equivalents as theta quotients, to give another system of forms appertaining to the same invariant set.

That the number of indeterminates is a quadratic polynomial in the given integer (m in the Theorem), follows from the multiplication of elliptic functions at once, and is obvious from the classic transcendental solution of the problem of real multiplication.

Finally, it is obvious that invariant sets arise from the real multiplication of *any* elliptic function. The like applies to the wider theory of transformation.

THE ARITHMETIC OF POLYNOMIALS IN A GALOIS FIELD.*

By LEONARD CARLITZ.†

INTRODUCTION.

In this paper we consider some of the arithmetic properties of polynomials in an indeterminate, x, with coefficients lying in some (finite) Galois field. The totality of such polynomials forms a ring in which arithmetic appears to be simpler than that of the ordinary integers; on the other hand, it is by no means a trivial arithmetic. What treatment of the subject appears in the literature does not seem to sufficiently stress the inherent simplicity of the polynomial domain: theorems are proved by utilizing the analogy with the rational domain. An instance of this is Artin's ‡ proof of Dedekind's § Law of Quadratic Reciprocity.

The first section of this paper is concerned with the treatment of the simplest numerical functions connected with the domain under consideration. An interesting feature is that the analogs of many difficult asymptotic formulas of ordinary arithmetic are here capable of simple explicit expression. A well-known case is the expression for the number of (primary) irreducible polynomials of a given degree. We give a number of others of the same type. For example, if $\tau(F)$ denote the number of primary divisors of the polynomial F, then

$$\sum \tau(F) = (\nu + 1) p^{\pi\nu},$$

the summation extending over the primary polynomials F of degree ν; again, if $Q(\nu)$ denote the number of primary "quadratfrei" polynomials of degree ν, then

$$Q(\nu) = p^{\pi\nu} - p^{\pi(\nu-1)}.$$

The method used in proving these and similar results is not elementary in that it depends on equating coefficients in equal power series. However it appears to be the natural one for the subject, and there seems to be little point to recasting the proofs in "arithmetic" shape.

* Presented to the American Mathematical Society, December 30, 1930.

† National Research Fellow.

‡ E. Artin, "Quadratische Körper im Gebiete der höheren Kongruenzen," *Mathematische Zeitschrift*, Vol. 19 (1924), pp. 153-246.

§ R. Dedekind, "Abriss einer Theorie der höheren Congruenzen in Bezug auf einer reellen Primzahl-Modulus," *Journal für die reine und angewandte Mathematik*, Vol. 54 (1857), pp. 1-26.

In the second section a theorem of reciprocity is set up which includes Dedekind's quadratic law as a special case. If the underlying Galois field be of order p^π, we define the residue character of index $p^\pi - 1$

$$\left\{ \frac{A}{P} \right\} \equiv A^{(p^{\pi\nu}-1)/(p^\pi-1)}, \bmod P;$$

then if P, Q are primary and irreducible of respective degrees ν, ρ, we prove that

$$\left\{ \frac{P}{Q} \right\} = (-1)^{\rho\nu} \left\{ \frac{Q}{P} \right\}.$$

From this we may of course deduce reciprocity of index any divisor of $p^\pi - 1$; on the other hand, to treat the case in which the index is not a divisor of $p^\pi - 1$ we have merely to sufficiently enlarge the polynomial domain by passing to a larger Galois field.

This theorem of reciprocity probably best indicates the comparatively simple nature of the ring of polynomials as opposed to ordinary arithmetic—where reciprocity of index higher than the second necessitates the consideration of algebraic fields. The proof here given is quite elementary and superficially resembles one of Kronecker's proofs of the ordinary quadratic reciprocity theorem.

1. Numerical Functions.

1. *Notation.* Take p any prime (including 2) and π any positive integer. Let $GF(p^\pi)$ denote the unique Galois field * of order p^π. If x be an indeterminate, $\mathfrak{D}(p^\pi, x)$ will denote the totality of polynomials in x with coefficients lying in the $GF(p^\pi)$. Evidently $\mathfrak{D}(p^\pi, x)$ is contained in $\mathfrak{D}(p^{\pi_1}, x)$ if and only if π_1 is a multiple of π. We shall find it convenient to employ the following notation. The elements (polynomials) of $\mathfrak{D}$ will be denoted by large Roman letters, rational integers will be denoted by small Greek and Roman letters, but small Roman a will be reserved for the elements of the Galois field. $A = B$ will of course mean that the coefficients of like powers of x are identical elements of the underlying Galois field. Further, if

$$F = a_0 x^\nu + a_1 x^{\nu-1} + \cdots + a_\nu,$$

then

$$\operatorname{sgn} F = a_0, \quad |F| = p^{\pi\nu};$$

F is *primary* if $\operatorname{sgn} F = 1$. If $F = aG$, then F and G are *associates*.

* The properties of Galois fields used here are developed in Dickson's *Linear Groups* (1901), pp. 3-54.

Evidently the number of primary polynomials of degree ν is $p^{\pi\nu}$. Hence the ζ-function in $\mathfrak{D}$,

$$(1) \qquad \zeta(s) = \sum_{\operatorname{sgn} F=1} 1/|F|^s = \sum_{\nu=0}^{\infty} p^{\pi\nu}/p^{\pi\nu s} = 1/[1-1/p^{\pi(s-1)}]$$

Also since decomposition of the elements of $\mathfrak{D}$ into irreducible polynomials is essentially unique we have the Euler factorization

$$(2) \qquad \zeta(s) = \sum 1/|F|^s = \Pi\, 1/(1-1/|P|^s),$$

the product extending over all primary *irreducible* polynomials P in $\mathfrak{D}$; both series and product converge absolutely for $s > 1$.

From (2) we may immediately derive the well-known expression for $\Psi(\nu)$, the number of (primary) irreducible polynomials of degree ν: Indeed * from (1) and (2) follows

$$\begin{aligned} 1-1/p^{\pi(s-1)} &= \prod_{\operatorname{sgn} P=1} (1-1/|P|^s) \\ &= \prod_{\nu=1}^{\infty} (1-1/p^{\pi\nu s})^{\psi(\nu)}, \end{aligned}$$

so that by taking logarithms

$$\sum_{\mu,\nu=1}^{\infty} \psi(\nu)/\mu p^{\pi\mu\nu s} = \sum_{\nu=1}^{\infty} 1/\nu p^{\pi\nu(s-1)},$$

whence †

$$\sum_{\delta\mid\nu} \delta\psi(\delta) = p^{\pi\nu},$$

which is equivalent to

$$(3) \qquad \psi(\nu) = (1/\nu)\,(p^{\pi\nu} - \sum p^{\pi\nu/\omega_1} + \sum p^{\pi\nu/\omega_1\omega_2} - \cdots),$$

$\omega_1, \omega_2, \cdots$, denoting distinct primes dividing ν.

We shall suppose in the following that all summations or products over a subset $\mathfrak{S}$ of $\mathfrak{D}$ extend only over *primary* polynomials in $\mathfrak{S}$ unless the contrary be stated.

2. *Definitions.* As numerical functions of primary interest we define $\mu, \lambda, \tau, \sigma, \phi, Q$.

(i) $\mu(F)$, the analog of the Möbius-μ, is defined by

---

* This proof for $\pi = 1$ is given by Landau in an editorial note on a paper by H. Kornblum, "Ueber die Primfunktionen in einer arithmetischen Progression," *Mathematische Zeitschrift*, Vol. 5 (1919), p. 107.

† For $\delta \mid \nu$ read, as usual, δ divides ν.

$$\begin{cases} \mu(a)=1,\ \mu(aF)=\mu(F), \\ \mu(F)=0 \text{ for } P^2 \mid F, \\ \mu(F)=(-1)^\rho \text{ for } F=aP_1 \cdots P_\rho\ (P_i \neq P_j). \end{cases}$$

(ii) $\lambda(F)$, the analog of the Liouville function, is defined by

$$\lambda(a)=1, \quad \lambda(aF)=\lambda(F),$$
$$\lambda(F)=(-1)^{a_1+\cdots+a_\rho} \quad \text{for} \quad F=aP_1^{a_1}\cdots P_\rho^{a_\rho}.$$

(iii) If $\operatorname{sgn} F=1$, $\tau(F)$ is the number of primary divisors of F:

$$\tau(F)=\sum_{D\mid F} 1,$$

extending over primary D only. More generally, $\tau_\rho(F)$ is the number of solutions of $F=D_1D_2\cdots D_{\rho+1}$ ($\operatorname{sgn} D_i=1$), so that $\tau(F)=\tau_1(F)$.

(iv) $$\sigma(F)=\sum_{D\mid F}\mid D\mid;$$

$\sigma_\rho(F)=\sum_{D\mid F}\mid D\mid^\rho$, so that $\sigma(F)=\sigma_1(F)$.

(v) After Dedekind, $\phi(F)$ is the number of elements in a reduced residue system, mod F. We define $\phi_\nu(F)$ to be the number of primary polynomials of degree ν that are prime to F, whence it is easy to see that

$$\phi_\nu(F)=\phi(F) \quad \text{if} \quad \deg F=\nu.$$

(vi) $Q(\nu)$ will denote the number of primary polynomials of degree ν not divisible by the square of an irreducible polynomial. More generally $Q_\rho(\nu)$ will denote the number of primary polynomials of degree ν not divisible by the $(\rho+1)$-th power of an irreducible polynomial; $Q_1(\nu)=Q(\nu)$.

3. *The λ- and μ-functions.* We note first that, exactly as in the rational case, μ appears in inversions of sums (or products) extended over the divisors of a fixed element. If $f(F)$ be any function in $\mathfrak{D}$ then ($\operatorname{sgn} F=1$)

(4) $$f'(F)=\sum_{D\mid F} f(D) \quad \text{and} \quad f(F)=\sum_{D\mid F}\mu(D)f'\left(\frac{F}{D}\right)$$

are equivalent. This follows at once from the identity

$$\sum f'(F)/\mid F\mid^s=\sum f(F)/\mid F\mid^s\cdot\sum 1/\mid F\mid^s.$$

However we may derive more interesting relations for both μ and λ—relations analogous to the asymptotic formulas connected with these functions in the rational domain. Since

$$\sum_F \mu(F)/\mid F\mid^s=\prod_P(1-1/\mid P\mid^s)=1/\zeta(s)=1-p^\pi/p^{\pi s},$$

we have immediately

$$(5)\qquad \begin{cases} \sum\limits_{\deg F=\nu} \mu(F) = 0 \quad \text{for} \quad \nu \geqq 2, \\ \sum\limits_{\deg F=1} \mu(F) = -p^{\pi}. \end{cases}$$

Similarly from

$$\sum_{F} \lambda(F)/|F|^{s} = \prod_{P} 1/(1+1/|P|^{s})$$
$$= \prod (1-1/|P|^{s})/(1-1/|P|^{2s}) = \zeta(2s)/\zeta(s)$$
$$= \sum_{\nu=0}^{\infty} p^{\pi\nu}/p^{2\pi\nu s}\,(1-p^{\pi}/p^{\pi s}),$$

$$(6)\qquad \sum_{\deg F=\nu} \lambda(F) = (-1)^{\nu}$$

where $[\alpha]$ is the greatest integer $\leqq \alpha$.

4. *The divisor functions.* Clearly from the definition of $\tau_\rho(F)$,

$$\sum_{F} \tau_{\rho}(F)/|F|^{s} = (\sum_{F} 1/|F|^{s})^{\rho+1} = (1-p^{\pi}/p^{\pi s})^{-\rho-1}$$
$$= \sum_{\nu=0}^{\infty} \binom{\nu+\rho}{\rho} p^{\pi\nu}/p^{\pi\nu s},$$

whence equating coefficients,

$$(7)\qquad \sum_{\deg F=\nu} \tau_{\rho}(F) = \binom{\nu+\rho}{\rho} p^{\pi\nu}.$$

The rational analog of this, it will be recalled, is an exceedingly complex asymptotic formula.

Again for $\sigma_\rho(F)$, from

$$\sum_{F} \sigma_{\rho}(F)/|F|^{s} = \sum 1/|F|^{s} \cdot \sum |F|^{\rho}/|F|^{s}$$
$$= \sum p^{\pi\nu}/p^{\pi\nu s} \cdot \sum p^{\pi\nu(\rho+1)}/p^{\pi\nu s}$$

we find without difficulty that

$$(8)\qquad \sum_{\deg F=\nu} \sigma_{\rho}(F) = p^{\pi\nu}(p^{\pi\rho(\nu+1)}-1)/(p^{\pi\rho}-1).$$

5. *The generalized ϕ-functions.* In order to evaluate $\phi_\nu(M)$ we start with

$$\sum_{(F,M)=1} 1/|F|^{s} = \sum_{\nu=0}^{\infty} \phi_{\nu}(M)/p^{\rho\pi\nu s},$$

where (F, M) is the "greatest" common divisor of F and M. But

$$\sum_{(F,M)=1} 1/|F|^s = \prod_{P|M} 1/(1-1/|P|^s) = \zeta(s) \prod_{P|M} (1-1/|P|^s)$$
$$= \sum_{\nu} p^{\pi\nu}/p^{\pi\nu s} \cdot \prod_{P|M} (1-1/|P|^s);$$

therefore

$$\phi_\nu(M) = p^{\pi\nu} \prod_{P|M} (1-1/|P|^s), \tag{9}$$

where in the right member we retain only those terms of the expanded product which are of non-negative degree in p^π. When $\nu = \deg M$ this coincides with Dedekind's * expression.

For this case it is convenient to rewrite (9) as

$$\phi(M) = \sum_{D|M} \mu\left(\frac{M}{D}\right) |D|;$$

then

$$\sum \phi(M)/|M|^s = \sum \mu(M)/|M|^s \sum |F|/|F|^s$$
$$= (1-p^\pi/p^{\pi s}) \sum_{\nu} p^{2\pi\nu}/p^{\pi\nu s},$$

so that

$$\sum_{\deg F=\nu} \phi(F) = p^{2\pi\nu} - p^{\pi(2\nu-1)}. \tag{10}$$

We pass over the immediate generalization of this to ϕ-functions of higher order.

6. *Formula for* $Q_\rho(\nu)$.

$$\sum_{\nu} Q_\rho(\nu)/p^{\pi\nu s} = \sum_{P^{\rho+1} \nmid F} 1/|F|^s = \prod_P (1+1/|P|^s + \cdots + 1/|P|^{\rho s})$$
$$= \prod [1-1/|P|^{(\rho+1)s}]/[1-1/|P|^s] = \zeta(s)/\zeta((\rho+1)s)$$
$$= \sum p^{\pi\nu}/p^{\pi\nu s}(1-p^\pi/p^{\pi s(\rho+1)}),$$

so that

$$\begin{cases} Q_\rho(\nu) = p^{\pi\nu} - p^{\pi(\nu-\rho)} & \text{for } \nu > \rho, \\ Q_\rho(\nu) = p^{\pi\nu} & \text{for } \nu \leqq \rho. \end{cases} \tag{ii}$$

7. *Other functions.* It would be possible to give an indefinite number of additional functions satisfying relations like (5), (7), (8), (10), (11). In general they are more complicated; everything depends of course on the particular combination of ζ's involved. For example defining $\vartheta(\nu)$ by

$$\sum \vartheta(\nu)/p^{\pi\nu s} = \prod (1+1/|P|^{2s} + 1/|P|^{3s} + \cdots)$$

—the arithmetic significance of ϑ is obvious—we find that

$$\sum \vartheta(\nu)/p^{\pi\nu s} = \zeta(2s)\zeta(3s)/\zeta(6s),$$

* *Loc. cit.*

and $\vartheta(\nu)$ is found as a rather involved function of ν. On the other hand, defining $\Theta(\nu)$ by

$$\sum \Theta(\nu)/p^{\pi\nu s} = \prod (1 + 1/|P|^{2s} - 1/|P|^{3s} + 1/(P)^{4s} - \cdots)$$
$$= \zeta(2s)/\zeta(3s) = \sum p^{\pi\nu}/p^{2\pi\nu s}(1 - p^{\pi}/p^{3\pi s}),$$

we see that

(12) $$\begin{cases} \Theta(\nu) = (-1)^{\nu} p^{\pi[\nu/2]} & \text{for } \nu \neq 1, \\ \Theta(1) = 0. \end{cases}$$

However ϑ is the simpler function arithmetically.

It is clear that generally if

$$\sum f(\nu)/p^{\pi\nu s} = \zeta(\alpha s)/[\zeta(\beta s) \cdots \zeta(\gamma s)],$$

we may deduce a simple explicit formula for f; and its arithmetic meaning is rather easily grasped. The next case of interest is

$$\sum f(\nu)/p^{\pi\nu s} = [\zeta(\alpha s) \cdots \zeta(\beta s)]/[\zeta(\gamma s) \cdots \zeta(\delta s)],$$

but it is unnecessary to consider this in detail here.

II. The Reciprocity Theorem.

8. *The Euler Criterion.* Consider the binomial congruence

(13) $$X^{p^{\pi}-1} \equiv A, \bmod P, \quad P \nmid A,$$

P as usual being irreducible and primary. Exactly as in the rational theory we may prove the

Euler Criterion. A necessary and sufficient condition that (13) *be solvable in* $\mathfrak{D}$ *is that*

$$\left\{ \frac{A}{P} \right\} = 1,$$

where $\{A/P\}$, *the* $(p^{\pi} - 1)$ *-ic power character, is defined by*

$$\left\{ \frac{A}{P} \right\} \equiv A^{(p^{\pi\nu}-1)/(p^{\pi}-1)}, \bmod P \qquad (\deg P = \nu).$$

Evidently

$$\left\{ \frac{AB}{P} \right\} = \left\{ \frac{A}{P} \right\} \left\{ \frac{B}{P} \right\},$$

so that the calculation of $\{A/P\}$ depends upon that of $\{Q/P\}$ and $\{a/P\}$, Q irreducible and primary. For $\{a/P\}$ we shall content ourselves with

$$\left\{ \frac{A}{P} \right\} = a^{\nu}, \tag{14}$$

the proof of which is immediate. For $\{Q/P\}$ we have the theorem of reciprocity stated in the Introduction:

$$\left\{ \frac{Q}{P} \right\} = (-1)^{\rho\nu} \left\{ \frac{Q}{P} \right\}, \quad \rho = \deg Q. \tag{15}$$

We shall prove this by establishing three lemmas, each of a simple nature.

9. *Analog of Gauss' Lemma.* Let $\mathfrak{R}(A/P)$ denote the remainder in the division of A by P. Then we have

LEMMA 1. *If H run through the (primary) polynomials of degree less than the degree of P,*

$$\left\{ \frac{A}{P} \right\} = \prod_H \operatorname{sgn} \mathfrak{R} \left\{ \frac{AH}{P} \right\}.$$

Since the proof is so much like the proof of Gauss's Lemma in the ordinary case we shall omit it.

10. LEMMA 2. *Let A be primary of degree $\alpha \geqq \nu$. Then if $P \nmid A$,*

$$\operatorname{sgn} \prod (A - KP) = (-1)^{\alpha-\nu} \operatorname{sgn} \mathfrak{R} \left\{ \frac{A}{P} \right\},$$

the product extending over all primary K's of degree $\alpha - \nu$.

Let K_0 be the quotient of A by P, and put

$$K = K_0 - M, \deg M < \alpha - \nu.$$

Then $A - KP = (A - K_0P) + MP$, so that if $K \neq K$,

$$\deg (A - KP) = \deg MP, \quad \operatorname{sgn} (A - KP) = \operatorname{sgn} MP = \operatorname{sgn} M.$$

We now divide the M's according to degree (excluding $M = 0$). Evidently there are precisely

$$p^{\pi(\gamma+1)} - p^{\pi\gamma}$$

of degree γ. Now by the extension of Wilson's Theorem the product of the non-zero elements of a $GF(p^{\pi}) = -1$. Hence it is easy to see that

$$\operatorname{sgn} \prod_{\deg M=\rho} M = -1,$$

so that

$$\operatorname{sgn} \prod_{K \neq K_0} (A - KP) = (-1)^{a-\nu};$$

and from this the Lemma follows immediately.

11. LEMMA 3. *If H runs through the primary polynomials of degree $< \nu$ and K those of degree $< \rho$,*

$$\operatorname{sgn} \prod_{H,K} (HQ - KP) = (-1)^{\rho\nu + \operatorname{Min}(\rho^2, \nu^2)} \operatorname{sgn} \prod_{H} \mathfrak{R} \left\{ \frac{HQ}{P} \right\}.$$

Evidently

$$(16) \qquad \begin{aligned} \operatorname{sgn} \prod_{H,K} (HQ - KP) &= \operatorname{sgn} \prod_{|HQ|>|KP|} \cdot \operatorname{sgn} \prod_{|HQ|=|KP|} \cdot \operatorname{sgn} \prod_{|HQ|<|KP|} \\ &= \operatorname{sgn} \prod_{|HQ|=|KP|} \cdot \operatorname{sgn} \prod_{|HQ|<|KP|} . \end{aligned}$$

We now consider separately (A). $\rho \geqq \nu$, (B). $\rho < \nu$.

(A). $\rho \geqq \nu$. In order to determine

$$\operatorname{sgn} \prod_{|HQ|=|KP|} (HQ - KP),$$

rewrite this as

$$\prod_{H} \operatorname{sgn} \prod_{K} (HQ - KP).$$

Letting h, k be the respective degrees of H, K, and noting that $h + \rho = k + \nu$, we may apply Lemma 2 to the inner product, which accordingly becomes

$$(-1)^{h+\rho-\nu} \operatorname{sgn} \mathfrak{R} \left\{ \frac{HQ}{P} \right\}.$$

Then the double product

$$(17) \qquad \begin{aligned} &= \prod_{H} (-1)^{h+\rho-\nu} \cdot \operatorname{sgn} \prod_{H} \mathfrak{R} \left(\frac{HQ}{P} \right) \\ &= \prod_{h=0}^{\nu-1} (-1)^{(h+\rho-\nu)p^{\pi h}} \cdot \operatorname{sgn} \prod \mathfrak{R} \left(\frac{HQ}{P} \right) \\ &= (-1)^{\frac{1}{2}\nu(\nu-1)+\rho\nu+\nu^2} \cdot \operatorname{sgn} \prod \mathfrak{R} \left(\frac{HQ}{P} \right). \end{aligned}$$

As for

$$\operatorname{sgn} \prod_{|HQ|<|KP|} (HQ - KP),$$

we see from $h + \rho < k + \nu$, that for

$$\begin{array}{ll} h = 0, & k = \rho - \nu + 1, \cdots, \rho - 1, \\ h = 1, & k = \rho - \nu + 2, \cdots, \rho - 1, \\ \cdot \quad \cdot \quad \cdot & \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ h = \nu - 2, & k = \rho - 1. \end{array}$$

Therefore

$$\operatorname{sgn} \prod_{|HQ|<|KP|} = \prod_{h=0}^{\nu-2} \left[\prod_{k=\rho-\nu+h+1}^{\rho-1} (-1)^{p^{\pi h}} \right]^{p^{\pi k}}$$

$$(18) \qquad = (-1)^{\frac{1}{2}\nu(\nu-1)}.$$

Substituting from (17) and (18) in (16),

$$(19) \qquad \operatorname{sgn} \prod_{H,K} = (-1)^{\rho\nu+\nu^2} \operatorname{sgn} \prod \mathcal{R}\left(\frac{HQ}{P}\right) \qquad (\rho \geqq \nu).$$

(B). $\rho < \nu$.

We now find that

$$\operatorname{sgn} \prod_{|HQ|=|KP|} = \prod_{H} \operatorname{sgn} \prod_{K} (HQ - KP)$$

$$= \prod_{h=\nu-\rho}^{\nu-1} (-1)^{h+\rho-\nu} \cdot \operatorname{sgn} \prod \mathcal{R}\left(\frac{HQ}{P}\right)$$

$$(20) \qquad = (-1)^{\frac{1}{2}\rho(\rho-1)} \cdot \operatorname{sgn} \prod \mathcal{R}\left(\frac{HQ}{P}\right).$$

For the second partial product in (16) we now notice that for

$$\begin{array}{ll}
h = 0, & k = 0, \cdots, \rho - 1, \\
\cdot \quad \cdot \quad \cdot \quad \cdot & \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\
h = \nu - \rho - 1, & k = 0, \cdots, \rho - 1, \\
h = \nu - \rho, & k = 1, \cdots, \rho - 1, \\
\cdot \quad \cdot \quad \cdot \quad \cdot & \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\
h = \nu - 2, & k = \rho - 1;
\end{array}$$

accordingly

$$\operatorname{sgn} \prod_{|HQ|<|KP|} = \prod_{h=0}^{\nu-\rho-1} (-1)^{\rho} \cdot \prod_{h=\nu-\rho}^{\nu-2} (-1)^{\nu-h-1}$$

$$(21) \qquad = (-1)^{\rho(\nu-\rho)+\frac{1}{2}\rho(\rho-1)}.$$

Substituting from (20) and (21) in (16),

$$(22) \qquad \operatorname{sgn} \prod_{H,K} = (-1)^{\rho\nu+\rho^2} \operatorname{sgn} \prod \mathcal{R}\left(\frac{HQ}{P}\right) \qquad (\rho < \nu).$$

But (19) and (22) imply the truth of Lemma 3.

12. From Lemma 1 and 3 it follows that

$$\left\{ \frac{Q}{P} \right\} = (-1)^{\rho\nu + \operatorname{Min}(\rho^2, \nu^2)} \operatorname{sgn} \prod_{H,K} (HQ - KP).$$

Interchanging P and Q,

$$\left\{\frac{P}{Q}\right\} = (-1)^{\rho\nu + \operatorname{Min}(\rho^2,\nu^2)} \operatorname{sgn} \prod_{H,K} (KP - HQ).$$

Immediately then

$$\left\{\frac{Q}{P}\right\} = (-1)^{\epsilon} \left\{\frac{P}{Q}\right\},$$

where

$$\epsilon = \frac{p^{\pi\nu}-1}{p^{\pi}-1} \quad \frac{p^{\pi\rho}-1}{p^{\pi}-1} \equiv \rho\nu, \quad \bmod 2,$$

thus completing the proof of the theorem of reciprocity.

In the proof of Lemma 2 and 3 we have implicitly supposed $p \neq 2$. However in that case the proofs are even simpler and, since in the $GF(2^{\pi})$, $+1$ and -1 are the same, we may here write (15) in the simpler form

$$(23) \qquad \left\{\frac{Q}{P}\right\} = \left\{\frac{P}{Q}\right\} \qquad (p = 2).$$

13. *Generalizations.* If $M = P_1 \cdots P_m$, the P's not necessarily being distinct, and $(A, M) = 1$, we define

$$\left\{\frac{A}{M}\right\} = \left\{\frac{A}{P_1}\right\} \cdots \left\{\frac{A}{P_m}\right\}.$$

Let now

$$N = Q_1 \cdots Q_n, \quad \deg P_i = \nu_i, \quad \deg Q_i = \rho_i;$$

then

$$\begin{aligned}\left\{\frac{M}{N}\right\} &= \prod \left\{\frac{Q_i}{P_j}\right\} \\ &= \prod (-1)^{\rho_i\nu_j} \left\{\frac{P_j}{Q_i}\right\} \\ &= (-1)^{(\rho_1+\cdots+\rho_n)(\nu_1+\cdots+\nu_m)} \left\{\frac{M}{N}\right\}.\end{aligned}$$

Putting $\deg N = \rho_1 + \cdots + \rho_m = \rho$, $\deg M = \nu$, we replace (15) by the more general

$$(24) \qquad \left\{\frac{M}{N}\right\} = (-1)^{\rho\nu} \left\{\frac{M}{N}\right\},$$

in which M and N are primary and prime to each other.

If we suppose $\pi = 1$, and define the *quadratic* character

$$\left\{\frac{A}{P}\right\}_2 = \left\{\frac{A}{P}\right\}^{(p-1)/2},$$

then (15) becomes

$$\left\{\frac{Q}{P}\right\}_2 = (-1)^{\rho\nu(p-1)/2}\left\{\frac{P}{Q}\right\}_2,$$

which is Dedekind's theorem of reciprocity. Similarly we may deduce a theorem of reciprocity of any index δ, a divisor of $p^\pi - 1$.

On the other hand, as remarked in the Introduction, we may also treat the case of index δ, where $\delta \nmid (p^\pi - 1)$. For assuming $p \nmid \delta$, we may find an integer π_1 that $\pi \mid \pi_1$ and $\delta \mid (p^{\pi_1} - 1)$, and then it is only necessary to extend the domain $\mathfrak{D}(p^\pi, x)$ to the larger domain $\mathfrak{D}(p^{\pi_1}, x)$. The assumption $p \nmid \delta$ is of no moment since

$$A \equiv A^{p^{\pi\nu}} \equiv B^p, \quad \text{mod } P;$$

that is, every polynomial is a p-ic residue, mod P.

California Institute of Technology,
Pasadena, California.

DIVISION ALGEBRAS ASSOCIATED WITH AN EQUATION WHOSE GROUP HAS FOUR GENERATORS.

By MINA S. REES.

1. *Introduction.* The construction of all division algebras is the outstanding problem in the theory of linear algebras. In the history of this problem, the procedure has been to examine first the necessary and sufficient conditions that the constructed algebra be associative; and second, the conditions that it be a division algebra. Since L. E. Dickson's announcement [1] in 1905 of his discovery of a system of non-commutative division algebras, he has published three papers examining associativity conditions for such algebras. In the first of these papers,[2] algebras associated with cyclic equations were discussed; in the second,[3] general results were obtained for the abelian non-cyclic case, and for that type of non-abelian case where the Galois group of the basic equation was a solvable group. This study was carried forward by Williamson,[4] in a paper in which detailed associativity conditions were established for the general case of a two generator group, and for a significant special case of the three generator group. In a third paper in 1930, Dickson [5] gave a simplification of his previous method by which he reached specific results for the general case of the two and three generator problems. The second line of investigation, the study of the conditions under which a given associative algebra is a division algebra, has been advanced by Dickson,[6] Wedderburn,[7] and Albert.[8] In addition to these two aspects of the problem, there is a third phase which is concerned with normal division algebras. It is known that any division algebra can be normalized by a suitable

1. *Bulletin of the American Mathematical Society*, Vol. 22 (1905-06), p. 442.

2. "Linear Associative Algebras and Abelian Equations," *Transactions of the American Mathematical Society*, Vol. 15 (1914), pp. 31-46.

3. "New Division Algebras," *Transactions of the American Mathematical Society*, Vol. 28 (1926), pp. 207-34.

4. "Associativity of Division Algebras," *Transactions of the American Mathematical Society*, Vol. 30 (1928), pp. 111-25.

5. "Construction of Division Algebras," *Transactions of the American Mathematical Society*, Vol. 32 (1930), pp. 319-34.

6. See notes (2) and (3).

7. "A Type of Primitive Algebra," *Transactions of the American Mathematical Society*, Vol. 15 (1914), pp. 162-166.

8. Abstracts in the *Bulletin of the American Mathematical Society*, Vol. 36 (1930), p. 45; p. 198; p. 804.

extension of the field, so that its order is the square of an integer,[9] and a third line of study has established that all division algebras of orders four,[10] nine,[11] or sixteen [12] are among those for which associativity conditions have been found.

The present paper continues the investigation of associativity conditions, by considering the algebras related to an equation whose Galois group has four independent generators. Using Dickson's most recent simplification of his previous method,[13] we set up necessary and sufficient conditions that these algebras be associative. Since the method is inductive, all the results which Dickson obtained for the two or three generator cases [14] will be presupposed. The notation of the present paper is the same as in Dickson's except that $v, y, z, w, \epsilon, \delta$ of his paper are replaced by $z_1, w_1, z_2, w_2, \delta_1, \delta_2$ respectively. Numbered theorems and formulas in square brackets refer to theorems and formulas in his paper.

2. *Algebra Σ derived for an equation whose group has four generators.* Let $f(\xi)=0$ be an equation of degree $pqsh$ which is irreducible in F and has the roots

$$\chi^{a_1}\{\psi^{a_2}[\phi^{a_3}(\theta^{a_4}(i))]\} \qquad (a_1<h,\ a_2<s,\ a_3<p,\ a_4<q), \tag{1}$$

where χ, ψ, ϕ, θ are polynomials in i with coefficients in F, and the superscripts denote iteratives, not exponents.

Throughout this paper, the brackets { } will mean that each of the included symbols is to be interpreted as a polynomial in the succeeding symbols, where the final argument i is suppressed. Thus $\psi\{\theta\phi\}$ means $\psi[\theta(\phi(i))]$.

Let the roots satisfy the following relations:

$$\theta^q=i, \quad \phi^p=\theta^c, \quad \psi^s=\phi^b\{\theta^a\}, \quad \chi^h=\psi^l\{\phi^d\theta^e\}, \tag{2}$$

$$\theta\{\phi\}=\phi\{\theta^x\}; \quad \theta\{\psi\}=\psi\{\phi^{z_1}\theta^{w_1}\}, \quad \phi\{\psi\}=\psi\{\phi^{z_2}\theta^{w_2}\}, \tag{3}$$

$$\theta\{\chi\}=\chi\{\psi^{u_1}\phi^{v_1}\theta^{y_1}\}, \quad \phi\{\chi\}=\chi\{\psi^{u_2}\phi^{v_2}\theta^{y_2}\}, \quad \psi\{\chi\}=\chi\{\psi^{u_3}\phi^{v_3}\theta^{y_3}\}, \tag{4}$$

where q, p, s, h are the least positive integers for which relations of the type (2) hold.

---

9. Dickson, *Algebren und ihre Zahlentheorie*, p. 138.

10. See note (9) above. *Loc. cit.*, p. 46.

11. Wedderburn, "On Division Algebras," *Transactions of the American Mathematical Society*, Vol. 22 (1921), pp. 129-35.

12. Albert, "A Determination of All Division Algebras in Sixteen Units," *Transactions of the American Mathematical Society*, Vol. 31 (1929), pp. 253-60.

13. See note (5), page 51.

14. *Ibid.*

Let Σ be the algebra over F having the basis $i^m j^n k^t o^f$ ($m \leqq pqsh-1$, $n \leqq q-1$, $t \leqq p-1$, $f \leqq s-1$) such that $f(i)=0$ and

$$(5) \qquad j^q = g(i), \qquad j^r P(i) = P(\theta^r) j^r;$$
$$(6) \qquad k^p = \beta(i) j^e, \quad k^r P(i) = P(\phi^r) k^r, \quad k^r j^t = [j^{(r)}]^t k^r,$$

where $j^{(r)}$ is defined by [16];

$$(7) \qquad \begin{aligned} o^s = \sigma(i) j^a k^b, \quad o^r P(i) &= P(\psi^r) o^r, \\ o^r j^t &= [j^{(r)}]^t o^r, \\ o^r k^t &= [k^{(r)}]^t o^r, \end{aligned}$$

where

$$J \equiv j' \equiv \delta_1 j^{w_1} k^{e_1}, \qquad K \equiv k' \equiv \delta_2 j^{w_2} k^{e_2}$$

and

$$j^{(r)} = \delta_1(\psi^{r-1})(J^{w_1}K^{e_1})^{(r-2)}; \qquad k^{(r)} = \delta_2(\psi^{r-1})(J^{w_2}K^{e_2})^{(r-2)}$$

are found by [64]. We assume g, β, and σ to be different from zero.

Σ may be obtained from [Theorem 5] by replacing F by the field F_1, derived by adjoining to F the elementary symmetric functions of all of the roots $\psi^{a_2}\{\phi^{a_3}\theta^{a_4}\}$. Then the latter are roots of an equation irreducible in F_1 of degree pqs.[15] Therefore by [Theorem 5], Σ has the above basis. Assume the conditions required in [Theorem 5] so that Σ be associative. We see that (5), (6), (7) are consequences of the associative law and the following set R of relations:

$$(8) \qquad j^q = g(i), \qquad ji = \theta(i) j, \qquad f(i) = 0,$$

$$(9) \qquad k^p = \beta(i) j^e, \qquad ki = \phi(i) k, \quad kj = \alpha(i) j^e k,$$

$$(10) \quad o^s = \sigma(i) j^a k^b, \quad oi = \psi(i) o, \quad oj = \delta_1(i) j^{w_1} k^{e_1} o, \quad ok = \delta_2(i) j^{w_2} k^{e_2} o.$$

These relations enable us to reduce any product AB to C of the same form, where A, B, and C are polynomials of Σ. Let R' denote the set of like relations, with i, j, k, o replaced by i', j', k', o' which are defined to be

$$(11) \qquad \begin{aligned} I = i' = \chi(i), \qquad J &= j' = \epsilon_1(i) j^{v_1} k^{v_1} o^{u_1}, \\ K &= k' = \epsilon_2(i) j^{v_2} k^{v_2} o^{u_2}, \\ O &= o' = \epsilon_3(i) j^{v_3} k^{v_3} o^{u_3}. \end{aligned}$$

If relations R' hold, it follows[16] that

$$C' = (AB)' = A'B', \qquad (A+B)' = A' + B'.$$

15. See note (3), page 51. *Loc. cit.*, § 6.

16. For further details see Dickson, "Construction of Division Algebras," *Transactions of the American Mathematical Society*, Vol. 32 (1930), p. 322.

Thus [Theorem 1] will enable us to construct an associative algebra Γ provided the correspondence defined in Σ is such that the following conditions are satisfied:

(a) Relations R' hold, so that

$$(AB)^{(r)} = A^{(r)}B^{(r)}, \qquad (A+B)^{(r)} = A^{(r)} + B^{(r)}.$$

(b) An element γ is defined for which $\gamma = \gamma'$.

(c) $A^{(h)}\gamma = \gamma A$ for every A in Σ.

3. *Associativity Conditions for* Γ. We investigate first the conditions for R' to be true. It is easily seen that $(8'_2)$, $(9'_2)$ and $(10'_2)$ are satisfied identically by reason of the definition of the correspondence in Σ. Insert the values (11) in $(8'_2)$. Then, by (4_1),

$$\begin{aligned} JI &= \epsilon_1\chi\,\{\psi^{u_1}\phi^{v_1}\theta^{y_1}\}\,j^{y_1}k^{v_1}o^{u_1}, \\ &= \epsilon_1\theta\{\chi\}j^{y_1}k^{v_1}o^{u_1} = \theta\{\chi\}J. \end{aligned}$$

Similarly, by using subscripts 2 and 3, relations $(9'_2)$ and $(10'_2)$ are established by (4_2) and (4_3) respectively.

Preliminary Definitions and Formulas. We shall use the following abbreviations, corresponding to [36], [38], [39] respectively:

$$(12) \qquad \begin{aligned} B_{m,n}(i) &\equiv \pi_n(\phi^{m-1})\pi_{xn}(\phi^{m-2})\pi_{x^2n}(\phi^{m-3})\cdots\pi_{x^{m-1}n}(i), \\ \pi_n(i) &= \alpha\alpha(\theta^x)\cdots\alpha(\theta^{(n-1)x}), \qquad B_{m,0} = B_{0,n} = 1, \end{aligned}$$

$$(13) \qquad h(t) \equiv \sum_{n=0}^{t-1} x^{ns_1}w_1, \qquad h(0) = 0,$$

$$(14) \qquad C_d(i) \equiv \prod_{n=0}^{d-1} \delta_1\{\phi^{ns_1}\theta^{h(n)}\}B_{ns_1,w_1}\{\theta^{h(n)}\}, \qquad C_0(i) \equiv 1,$$

and the formulas [35] and [37]

$$(15) \qquad k^mj^n = B_{m,n}j^{x^mn}k^m,$$

$$(16) \qquad \delta j^wk^s\epsilon j^yk^v = \delta\epsilon\{\phi^s\theta^w\}B_{s,\bar{y}}(\theta^w)\,j^{w+x^sy}k^{s+v},$$

where δ and ϵ are polynomials in i. Then from (10_3) we see by induction on n, that

$$(17) \qquad oj^n = C_n(i)\,j^{h(n)}k^{ns_1}o.$$

Similarly, if $H(n)$ is derived from $h(n)$ by using subscripts 2 throughout, and $D_d(i)$ from $C_d(i)$ by using subscripts 2 and substituting $H(n)$ for $h(n)$, we have

$$ok^n = D_n(i) j^{H(n)} k^{nz_2} o. \tag{18}$$

Formula [58] is

$$t(m, n) \equiv h(m) + x^{mz_1} H(n).$$

From (17) and (18), by induction on m, we can deduce formulas corresponding to (15). We find, in fact

$$\left\{\begin{aligned} o^m j^n &= \mathcal{A}_{mn}(i) j^{A(m,n)} k^{L(m,n)} o^m, \\ o^m k^n &= \mathcal{D}_{mn}(i) j^{R(m,n)} k^{S(m,n)} o^m, \end{aligned}\right. \tag{20}$$

where (21)-(24) hold. The recursion formulas

$$\left\{\begin{aligned} &A(1, n) = h(n), && A(m+1, n) = t\,[A(m, n),\ L(m, n)], \\ &L(1, n) = nz_1, && L(m+1, n) = A(m, n) z_1 + L(m, n) z_2, \end{aligned}\right. \tag{21}$$

$$\left\{\begin{aligned} &R(1, n) = H(n), && R(m+1, n) = t\,[R(m, n),\ S(m, n)], \\ &S(1, n) = nz_2, && S(m+1, n) = R(m, n) z_1 + S(m, n) z_2, \end{aligned}\right. \tag{22}$$

with

$$\begin{aligned} &A(0, n) = S(0, n) = n, \qquad L(0, n) = R(0, n) = 0, \\ &A(m, 0) = L(m, 0) = R(m, 0) = S(m, 0) = 0; \end{aligned}$$

and the abbreviations

$$F_{m,n}(i) = C_m D_n \{\phi^{mz_1} \theta^{h(m)}\} B_{mz_1, H(n)} \{\theta^{h(m)}\}, \tag{23}$$

$$\left\{\begin{aligned} \mathcal{A}_{mn}(i) &= \prod_{r=0}^{m-1} F_{A(r,n),\ L(r,n)} \{\psi^{m-1-r}\}, \\ \mathcal{D}_{mn}(i) &= \prod_{r=0}^{m-1} F_{R(r,n),\ S(r,n)} \{\psi^{m-1-r}\}, \end{aligned}\right. \tag{24}$$

with

$$\mathcal{A}_{1n}(i) = C_n(i), \qquad \mathcal{D}_{1n}(i) = D_n(i), \qquad \mathcal{A}_{0n}(i) \equiv \mathcal{D}_{0n}(i) \equiv 1,$$

actually define the functions that occur in (20).

Applying (7_2), (20), (16), we have

$$\begin{aligned} KJ &= \epsilon_2 j^{v_2} k^{v_2} o^{u_2} \epsilon_1 j^{v_1} k^{v_1} o^{u_1}, \\ &= \epsilon_2 \epsilon_1 \{\psi^{u_2} \phi^{v_2} \theta^{v_2}\}\, \mathcal{A}_{u_2 v_1} \{\phi^{v_2} \theta^{v_2}\}\, \mathcal{D}_{u_2 v_1} \{\phi^{v_2 + L(u_2, v_1)} \theta^{v_2 + x^{v_2} A(u_2, v_1)}\} \\ &\quad \times B_{v_2, A(u_2, v_1)} \{\theta^{v_2}\}\, B_{v_2 + L(u_2, v_1),\, R(u_2, v_1)} \{\theta^{v_2 + x^{v_2} A(u_2, v_1)}\} \\ &\quad \times j^{v_2 + x^{v_2} A(u_2, v_1) + x^{v_2 + L(u_2, v_1)} R(u_2, v_1)} k^{v_2 + L(u_2, v_1) + S(u_2, v_1)} o^{u_2 + u_1}. \end{aligned} \tag{25_1}$$

Define

$$\begin{aligned} \mathcal{K}_{ab}(r, t) &\equiv v_a + \sum_{n=1}^{r-1} L(nu_a, y_b) + \sum_{m=1}^{t-1} S(mu_a, v_b), \qquad \mathcal{K}_{ab}(0, 0) = 0, \\ \mathcal{Q}_{ab}(r, t) &\equiv y_a + \sum_{n=1}^{r-1} x^{\mathcal{K}_{ab}(n,n)} A(nu_a, y_b) \\ &\qquad + \sum_{m=1}^{t-1} x^{\mathcal{K}_{ab}(m+1,m)} R(mu_a, v_b), \qquad \mathcal{Q}_{ab}(0, 0) = 0, \end{aligned} \tag{26}$$

where summations from one to zero are equal to zero;

$$(27) \qquad \mathcal{K}_{ab}(2,2) \equiv \mathcal{F}_{ab}, \qquad \mathcal{Q}_{ab}(2,2) \equiv \quad {}_{ab},$$

$$\begin{aligned}(28) \quad \mathcal{H}_{ab}(i) \equiv \epsilon_a\epsilon_b\{\psi^{u_a}\phi^{v_a}\theta^{y_a}\}\, &\mathcal{A}_{u_a y_b}\{\phi^{v_a}\theta^{y_a}\} \\ &\times \mathcal{D}_{u_a v_b}\{\phi^{\mathcal{K}_{ab}(2,1)}\theta^{\mathcal{Q}_{ab}(2,1)}\} B_{v_a, A(u_a, y_b)}\{\theta^{y_a}\} \\ &\times B_{\mathcal{K}_{ab}(2,1),\, R(u_a, v_b)}\{\theta^{\mathcal{Q}_{ab}(2,1)}\};\end{aligned}$$

$$\begin{aligned}(29) \quad \mathcal{G}^{(d)}(i) \equiv \prod_{n=0}^{d-1} \epsilon\{\psi^{nu}\phi^{\mathcal{K}(n,n)}\theta^{\mathcal{Q}(n,n)}\}\, &\mathcal{A}_{nu,y}\{\phi^{\mathcal{K}(n,n)}\theta^{\mathcal{Q}(n,n)}\} \\ &\times B_{\mathcal{K}(n,n),\, A(nu,y)}\{\theta^{\mathcal{Q}(n,n)}\} B_{\mathcal{K}(n+1,n),\, R(nu,v)}\{\theta^{\mathcal{Q}(n+1,n)}\} \\ &\times \mathcal{D}_{nu,v}\{\phi^{\mathcal{K}(n+1,n)}\theta^{\mathcal{Q}(n+1,n)}\} \qquad\qquad (d > 0), \\ \mathcal{G}^{(0)}(i) \equiv 1, \qquad \mathcal{G}^{(1)}(i) = \epsilon,&\end{aligned}$$

where a subscript on the $\mathcal{G}$ is to be interpreted as the subscript of every ϵ, u, v, y, and the double subscript of every $\mathcal{K}$, $\mathcal{Q}$. Using these abbreviations we obtain a generalized statement of (25_1)

$$(25_2) \qquad \epsilon_a j^{y_a} k^{v_a} o^{u_a} \epsilon_b j^{y_b} k^{v_b} o^{u_b} = \mathcal{H}_{ab}(i)\, j^{\mathcal{E}_{ab}} k^{\mathcal{F}_{ab}} o^{u_a+u_b}.$$

We have immediately

$$(30) \qquad \begin{aligned} KJ &= \mathcal{H}_{21}(i)\, j^{\mathcal{E}_{21}} k^{\mathcal{F}_{21}} o^{u_2+u_1}, \\ OJ &= \mathcal{H}_{31}(i)\, j^{\mathcal{E}_{31}} k^{\mathcal{F}_{31}} o^{u_3+u_1}, \\ OK &= \mathcal{H}_{32}(i)\, j^{\mathcal{E}_{32}} k^{\mathcal{F}_{32}} o^{u_3+u_2}. \end{aligned}$$

By induction on d, we obtain from (11_2) and (25_1)

$$(31) \qquad J^d = \mathcal{G}_1^{(d)} j^{\mathcal{Q}_{11}(d,d)} k^{\mathcal{K}_{11}(d,d)} o^{u_1 d}.$$

K^d and O^d are obtained from J^d by using subscripts 2 and 3 respectively.

In the discussion which follows, the numbers in the left-hand column of the table given below will be used as superscripts on $\mathcal{H}$, $\mathcal{E}$, $\mathcal{F}$, to indicate the results obtained from formulas (27) and (28) by making the tabulated substitutions. In (7) and (8), subscripts (n, t) and (m, n, t) respectively, will be used on $\mathcal{H}$, $\mathcal{E}$, $\mathcal{F}$. The last line of the table gives superscripts (9), (10), (11) when subscripts 1, 2, 3 respectively are used on y, v, and u.

	Superscript	ϵ_a	y_a	v_a	u_a	ϵ_b	y_b	v_b	u_b
	(1)	$\mathcal{G}_1^{(x)}$	$\mathcal{Q}_{11}(x, x)$	$\mathcal{K}_{11}(x, x)$	$u_1 x$	ϵ_2	y_2	v_2	u_2
	(2)	$\mathcal{G}_2^{(z_1)}$	$\mathcal{Q}_{22}(z_1 z_1)$	$\mathcal{K}_{22}(z_1, z_1)$	$u_2 z_1$	ϵ_3	y_3	v_3	u_3
	(3)	$\mathcal{G}_1^{(w_1)}$	$\mathcal{Q}_{11}(w_1, w_1)$	$\mathcal{K}_{11}(w_1, w_1)$	$u_1 w_1$	$\mathcal{H}^{(2)}$	$\mathcal{E}^{(2)}$	$\mathcal{F}^{(2)}$	$u_2 z_1 + u_3$
	(4)	$\mathcal{G}_2^{(z_2)}$	$\mathcal{Q}_{22}(z_2, z_2)$	$\mathcal{K}_{22}(z_2, z_2)$	$u_2 z_2$	ϵ_3	y_3	v_3	u_3
(32)	(5)	$\mathcal{G}_1^{(w_2)}$	$\mathcal{Q}_{11}(w_2, w_2)$	$\mathcal{K}_{11}(w_2, w_2)$	$u_1 w_2$	$\mathcal{H}^{(4)}$	$\mathcal{E}^{(4)}$	$\mathcal{F}^{(4)}$	$u_2 z_2 + u_3$
	(6)	$\mathcal{G}_1^{(a)}$	$\mathcal{Q}_{11}(a, a)$	$\mathcal{K}_{11}(a, a)$	$u_1 a$	$\mathcal{G}_2^{(b)}$	$\mathcal{Q}_{22}(b, b)$	$\mathcal{K}_{22}(b, b)$	$u_2 b$
	(7) Subscript (n, t)	$\mathcal{G}_2^{(n)}$	$\mathcal{Q}_{22}(n, n)$	$\mathcal{K}_{22}(n, n)$	$u_2 n$	$\mathcal{G}_3^{(t)}$	$\mathcal{Q}_{33}(t, t)$	$\mathcal{K}_{33}(t, t)$	$u_3 t$
	(8) Subscript (m, n, t)	$\mathcal{G}_1^{(m)}$	$\mathcal{Q}_{11}(m, m)$	$\mathcal{K}_{11}(m, m)$	$u_1 m$	$\mathcal{H}^{(7)}{}_{(n,t)}$	$\mathcal{E}^{(7)}{}_{(n,t)}$	$\mathcal{F}^{(7)}{}_{(n,t)}$	$u_2 n + u_3 t$
	(9), (10), or (11)	$\Omega_{y,v,u,k-2}$	η	λ	τ	ρ	c	d	l

where the arguments of η, λ, τ, in (9), (10), (11) are $(y, v, u, h-2)$. We shall need the formula

(33) $$k^{np}=\beta_n(i)j^{en}, \quad \beta_n(i)=\beta\beta(\theta^e)\cdots\beta(\theta^{e(n-1)}),$$

which is [42] and the formula

(34) $$o^{ns}=\sigma_n(i)j^{r(n)}k^{bn},$$

where

$$r(n)=\sum_{t=0}^{n-1}x^{tb}a, \qquad r(0)=0,$$

$$\sigma_n(i)=\prod_{t=0}^{n-1}\sigma\{\phi^{tb}\theta^{r(t)}\}B_{tb,a}\{\theta^{r(t)}\}, \qquad \sigma_0(i)\equiv 1.$$

which is obtained by induction on s, using (7_1) and (16).

In finding the associativity conditions for our algebra, we must discuss ten conditions of the type

(35) $$P(i)j^Uk^Vo^W=S(i)j^Xk^Yo^Z.$$

Since s is the least power of o which gives a polynomial in i, j, k, we have

(36) $$Z-W=c_1s, \qquad c_1 \text{ an integer.}$$

Replacing o^{c_1s} by its value (34), and applying (16), we have

$$\begin{aligned}P(i)j^Uk^V&=S(i)j^Xk^Y\sigma_{c_1}(i)j^{r(c_1)}k^{bc_1},\\ &=S(i)\sigma_{c_1}\{\phi^Y\theta^X\}B_{Y,r(c_1)}\{\theta^X\}j^{X+xYr(c_1)}k^{Y+bc_1}.\end{aligned}$$

By comparing exponents of k, since p is the least power of k which gives a polynomial in i and j, we see that

(37) $$Y+bc_1-V=c_2p, \qquad c_2 \text{ an integer.}$$

Replacing k^{c_2p} by its value (33), applying (5_2), and comparing exponents of j, we have

(38) $$X+x^Yr(c_1)+ec_2-U=c_3q, \qquad c_3 \text{ an integer.}$$

Replacing j^{c_3q} by its value from (5_1), we see that (35) is equivalent to the associativity condition

(39) $$P(i)=S(i)\sigma_{c_1}\{\phi^Y\theta^X\}B_{Y,r(c_1)}\{\theta^X\}\beta_{c_2}\{\theta^{X+xYr(c_1)}\}g^{c_3}.$$

In applying these results to special cases, identify the polynomials $P(i)$ and $S(i)$ with the given coefficients, and the exponents U, V, W, X, Y, Z, with

the given exponents in such a way that the c_1 determined by (36) is positive. Then it is formally possible that c_2 determined by (37) is negative. In this case we would have

$$V - Y - bc_1 = -c_2 p.$$

Proceeding as above, replacing $k^{-c_2 p}$ by its value (33) and comparing exponents of k, we find

$$X + x^Y r(c_1) - U + ec_2 = c_3 q$$

which agrees with (38). The associativity condition for this case is

$$(40) \qquad P(i)\beta_{-c_2}\{\theta^U\} = S(i)\sigma_{c_1}\{\phi^Y\theta^X\}B_{Y,r(c_1)}\{\theta^X\}g^{c_3}.$$

We observe, therefore, that if we define

$$(41) \qquad \beta_{-n} \equiv [\beta_n\{\theta^{-en}\}]^{-1}$$

relation (40) follows from (39) since

$$X + x^Y r(c_1) + ec_2 \equiv U \pmod{q}, \qquad \theta^q = i.$$

We may now apply (36), (37), (38), (39), (41) as formulas to determine the desired associativity conditions. The relations corresponding to (35) will be stated, and the evaluation of the formulas left for special cases. Thus the associativity conditions for our algebra will be derivable from relations (42), (43), (44), (45), (46), (47), (50), (59), (60), (61) given below, by the application of formulas (39), (41).

We will discuss first $(9_3')$, $(10_3')$, $(10_4')$, $(8_1')$, $(9_1')$, $(10_1')$.

Condition that $KJ = \alpha(\chi)J^x K$. By (30_1), (31) and (32), this reduces to

$$\mathcal{H}_{21}(i)j^{\mathcal{E}_{21}}k^{\mathcal{F}_{21}}o^{u_2+u_1} = \alpha(\chi)\mathcal{G}_1^{(x)}(i)j^{\mathcal{L}_{11}(x,x)}k^{\mathcal{K}_{11}(x,x)}o^{u_1 x}\epsilon_2(i)j^{v_2}k^{v_2}o^{u_2}.$$

Hence

$$(42) \qquad \mathcal{H}_{21}(i)j^{\mathcal{E}_{21}}k^{\mathcal{F}_{21}}o^{u_2+u_1} = \alpha(\chi)\mathcal{H}^{(1)}(i)j^{\mathcal{E}^{(1)}}k^{\mathcal{F}^{(1)}}o^{u_1 x+u_2}.$$

Condition that $OJ = \epsilon_1(\chi)J^{w_1}K^{z_1}O$ (*subscripts* 1 *on* w *and* z).

$$\mathcal{H}_{31}(i)j^{\mathcal{E}_{31}}k^{\mathcal{F}_{31}}o^{u_3+u_1} = \epsilon_1(\chi)\mathcal{G}_1^{(w)}j^{\mathcal{L}_{11}(w,w)}k^{\mathcal{K}_{11}(w,w)}o^{u_1 w}$$
$$\times \mathcal{G}_2^{(z)}j^{\mathcal{L}_{22}(z,z)}k^{\mathcal{K}_{22}(z,z)}o^{u_2 z}\epsilon_3 j^{v_3}k^{v_3}o^{u_3}.$$

Hence

$$(43) \qquad \mathcal{H}_{31}(i)j^{\mathcal{E}_{31}}k^{\mathcal{F}_{31}}o^{u_3+u_1} = \epsilon_1(\chi)\mathcal{H}^{(3)}j^{\mathcal{E}^{(3)}}k^{\mathcal{F}^{(3)}}o^{u_1 w_1+u_2 z_1+u_3}.$$

Condition that $OK = \epsilon_2(\chi)J^{w_2}K^{z_2}O$ (*subscripts* 2 *on* w *and* z).

(44) $$\mathcal{H}_{32}(i) j^{\mathcal{E}_{32}} k^{\mathcal{F}_{32}} o^{u_3+u_2} = \epsilon_2(\chi)\, \mathcal{H}^{(5)} j^{\mathcal{E}^{(5)}} k^{\mathcal{F}^{(5)}} o^{u_1 w_2 + u_2 z_2 + u_3}.$$

Condition that $J^q = g(\chi)$.

(45) $$\mathcal{G}_1{}^{(q)} j^{\mathfrak{Q}_{11}(q,q)} k^{\mathcal{K}_{11}(q,q)} o^{u_1 q} = g(\chi).$$

Condition that $K^p = \beta(\chi) J^e$.

(46) $$\mathcal{G}_2{}^{(p)} j^{\mathfrak{Q}_{22}(p,p)} k^{\mathcal{K}_{22}(p,p)} o^{u_2 p} = \beta(\chi)\, \mathcal{G}_1{}^{(e)} j^{\mathfrak{Q}_{11}(e,e)} k^{\mathcal{K}_{11}(e,e)} o^{u_1 e}.$$

Condition that $O^s = \sigma(\chi) J^a K^b$.

$$\mathcal{G}_3{}^{(s)} j^{\mathfrak{Q}_{33}(s,s)} k^{\mathcal{K}_{33}(s,s)} o^{u_3 s} = \sigma(\chi) \mathcal{G}_1{}^{(a)} j^{\mathfrak{Q}_{11}(a,a)} k^{\mathcal{K}_{11}(a,a)} o^{u_1 a}$$
$$\mathcal{G}_2{}^{(b)} j^{\mathfrak{Q}_{22}(b,b)} k^{\mathcal{K}_{22}(b,b)} o^{u_2 b}.$$

Hence

(47) $$\mathcal{G}_3{}^{(s)} j^{\mathfrak{Q}_{33}(s,s)} k^{\mathcal{K}_{33}(s,s)} o^{u_3 s} = \sigma(\chi)\, \mathcal{H}^{(6)} j^{\mathcal{E}^{(6)}} k^{\mathcal{F}^{(6)}} o^{u_1 a + u_2 b}.$$

We have now given all the conditions derived from (8′), (9′), (10′), in the form (35). Choose as γ,

(48) $$\gamma = \rho(i) j^c k^d o^l, \qquad [\rho(i) \neq o].$$

Using (25), (31), (32), we establish the formula

(49) $$J^m K^n O^t = \mathcal{H}^{(8)}_{(m,n,t)} j^{\mathcal{E}^{(8)}_{(m,n,t)}} k^{\mathcal{F}^{(8)}_{(m,n,t)}} o^{u_1 m + u_2 n + u_3 t}.$$

Condition that $\gamma = \gamma'$ [*subscripts* (c, d, l) on $\mathcal{H}$, $\mathcal{E}$, $\mathcal{F}$].

$$\rho(i) j^c k^d o^l = \rho(\chi) J^c K^d O^l,$$

(50) $$\rho(i) j^c k^d o^l = \rho(\chi) \mathcal{H}^{(8)} j^{\mathcal{E}^{(8)}} k^{\mathcal{F}^{(8)}} o^{u_1 c + u_2 d + u_3 l}.$$

In order to construct our associative algebra Γ, we need the further condition that

(51) $$A^{(h)}\gamma = \gamma A$$

for all A in Σ. Evidently this condition holds for $A = P(i)$, by reason of our definition of γ. For, by (5), (6), (7), and (2_4),

$$\gamma P(i) = \rho P\{\psi^l \phi^d \theta^c\} j^c k^d o^l,$$
$$= P(\chi^h)\gamma = P^{(h)}\gamma.$$

We must now find the condition that (51) holds when A takes the values j, k, o, respectively.

Condition that $j^{(h)}\gamma = \gamma j$ (*subscripts* 1 *on every* ϵ, u, v, y).

$$j' = \epsilon j^y k^v o^u, \qquad j'' = \epsilon(\chi) J^y K^v O^u.$$

We need a formula of the type

(52) $$(J^m K^n O^t)^{(r)} = \Omega_{(m,n,t,r)}(i) j^{\eta(m,n,t,r)} k^{\lambda(m,n,t,r)} o^{\tau(m,n,t,r)}.$$

Hence

(53) $$\Omega_{(m,n,t,r)}(\chi) J^{\eta(m,n,t,r)} K^{\lambda(m,n,t,r)} O^{\tau(m,n,t,r)}$$

must be identically equal to

(54) $$\Omega_{(m,n,t,r+1)}(i)\, j^{\eta(m,n,t,r+1)}\, k^{\lambda(m,n,t,r+1)}\, o^{\tau(m,n,t,r+1)}.$$

But by (49), (53) reduces to

$$\Omega_{(m,n,t,r)}(\chi)\, \mathcal{H}^{(8)}_{(\eta,\lambda,\tau)}\ (i) j^{\mathcal{E}^{(8)}_{(\eta,\lambda,\tau)}} k^{\mathcal{F}^{(8)}_{(\eta,\lambda,\tau)}} o^{u_1\eta+u_2\lambda+u_3\tau},$$

where the arguments of η, λ, τ are (m, n, t, r). Therefore we have the recursion formulas

(55) $$\begin{aligned} \tau(m,n,t,r+1) &= u_1\eta + u_2\lambda + u_3\tau, \\ \lambda(m,n,t,r+1) &= \mathcal{F}^{(8)}_{(\eta,\lambda,\tau)}, \\ \eta(m,n,t,r+1) &= \mathcal{E}^{(8)}_{(\eta,\lambda,\tau)}, \\ \Omega_{(m,n,t,r+1)}(i) &= \Omega_{(m,n,t,r)}(\chi)\, \mathcal{H}^{(8)}_{(\eta,\lambda,\tau)}, \end{aligned}$$

where η, λ, τ have arguments (m, n, t, r). Since we may identify (71) for $r = 0$ with (60), we have the initial values

(56) $$\begin{aligned} \tau(m,n,t,o) &= u_1 m + u_2 n + u_3 t, \\ \lambda(m,n,t,o) &= \mathcal{F}^{(8)}_{(m,n,t)}, \\ \eta(m,n,t,o) &= \mathcal{E}^{(8)}_{(m,n,t)}, \\ \Omega_{(m,n,t,o)}(i) &= \mathcal{H}^{(8)}_{(m,n,t)}\ (i). \end{aligned}$$

For these initial values, (55) and (56) serve as formulas to determine τ, λ, η, Ω completely. Now

(57) $$\begin{aligned} j^{(h)}\gamma &= [\epsilon(\chi) J^y K^v O^u]^{(h-2)} \rho(i) j^c k^d o^l, \\ &= \epsilon(\chi^{h-1})\, \mathcal{H}^{(9)} j^{\mathcal{E}^{(9)}} k^{\mathcal{F}^{(9)}} o^{\tau(y,v,u,h-2)+l}. \end{aligned}$$

where every ϵ, y, v, u, has subscript 1. By 24_1

(58) $$\begin{aligned} \gamma j &= \rho j^c k^d o^l j \\ &= \rho \mathfrak{A}_{l1}\{\phi^d\theta^c\} B_{d,A(l,1)}\{\theta^c\} j^{c+x^d A(l,1)} k^{d+L(l,1)} o^l. \end{aligned}$$

The condition that $j^{(h)}\gamma = \gamma j$ is therefore equivalent to

(59) $$\begin{aligned} &\epsilon_1(\chi^{h-1})\, \mathcal{H}^{(9)} j^{\mathcal{E}^{(9)}} k^{\mathcal{F}^{(9)}} o^{\tau(y_1,v_1,u_1,h-2)+l} \\ &\qquad = \rho \mathfrak{A}_{l1}\{\phi^d\theta^c\}\, B_{d,A(l,1)}\{\theta^c\} j^{c+x^d A(l,1)} k^{d+L(l,1)} o^l. \end{aligned}$$

Condition that $k^{(h)}\gamma = \gamma k$. In (59), replace

$$\mathcal{H}^{(9)}, \mathcal{E}^{(9)}, \mathcal{F}^{(9)} \text{ by } \mathcal{H}^{(10)}, \mathcal{E}^{(10)}, \mathcal{F}^{(10)},$$

$$\epsilon_1, y_1, v_1, u_1 \text{ by } \epsilon_2, y_2, v_2, u_2,$$

$$\mathcal{a}, A, L \text{ by } \mathcal{D}, R, S.$$

The condition is

$$(60) \qquad \epsilon_2(\chi^{h-1})\mathcal{H}^{(10)} j^{\mathcal{E}^{(10)}} k^{\mathcal{F}^{(10)}} o^{\tau(y_2, v_2, u_2, h-2)+l} = \rho \mathcal{D}_{l1}\{\phi^d\theta^o\} B_{d,R(l,1)}\{\theta^o\} j^{o+xdR(l,1)} k^{d+S(l,1)} o^l.$$

Condition that $o^{(h)}\gamma = \gamma o$. In the left member of (59), replace

$$\mathcal{H}^{(9)}, \mathcal{E}^{(9)}, \mathcal{F}^{(9)} \text{ by } \mathcal{H}^{(11)}, \mathcal{E}^{(11)}, \mathcal{F}^{(11)},$$

$$\epsilon_1, y_1, v_1, u_1 \text{ by } \epsilon_3, y_3, v_3, u_3.$$

The condition is

$$(61) \qquad \epsilon_3(\chi^{h-1})\mathcal{H}^{(11)} j^{\mathcal{E}^{(11)}} k^{\mathcal{F}^{(11)}} o^{\tau(y_3, v_3, u_3, h-2)+l} = \rho(i) j^o k^d o^{l+1}.$$

By [Theorem 1] with p replaced by h, we now have the following

THEOREM. *Let an equation, $f(\xi)=0$, be of degree pqsh, be irreducible in a field F, and have the roots (1) satisfying relations (2), (3), (4).* Consider the algebra over F having the basis $i^m j^n k^t o^f E^r$ $(m \leqq pqsh-1,$ $n \leqq q-1,$ $t \leqq p-1,$ $f \leqq s-1,$ $r \leqq h-1)$ with $f(i)=0$, (5), (6), (7), and

$$E^h = \rho(i) j^o k^d o^l, \qquad \begin{aligned} E^r P(i) &= P(\chi^r) E^r, \\ E^r j^t &= [j^{(r)}]^t E^r, \\ E^r k^t &= [k^{(r)}]^t E^r, \\ E^r o^t &= [o^{(r)}]^t E^r, \end{aligned}$$

where $j^{(r)}, k^{(r)}, o^{(r)}$ are equal to $\epsilon(\chi^{r-1})(J^y K^v O^u)^{(r-2)}$ with subscripts 1, 2, 3, respectively on ϵ, y, v, u, where J, K, O are as defined in (11). The values of $j^{(r)}, k^{(r)}, o^{(r)}$ are found by (52), (55), (56) while their t-th powers may be found by (26), (29), (31) with altered parameters. This algebra is associative if and only if conditions [6], [14], [15], [18], [45], [49], [54], [61], [69], [72] and the conditions derived from (42), (43), (44), (45), (46), (47), (50), (59), (60), (61) by the application of formulas (39), (41) hold.

We note that these twenty associativity conditions are consistent, since they involve only products of the parameters; thus they are satisfied when all the parameters are equal to unity.

We do not include as conditions the facts that the constants c_1, c_2, c_3,

obtained in each case by the application of formulas (36), (37), (38), are integers, since these are conditions for the existence of the group.[17]

4. *Example. A group of order* 32.

Consider the group of order 32 defined abstractly as follows:[18]

$$\Theta^4 = \Phi^2 = \Psi^2 = X^2 = 1,$$

$$\Theta\Phi = \Phi\Theta, \qquad \Theta X = X\Phi\Theta^3,$$

$$\Theta\Psi = \Psi\Phi\Theta, \qquad \Phi X = X\Phi,$$

$$\Phi\Psi = \Psi\Phi, \qquad \Psi X = X\Psi.$$

This group has an abelian subgroup of order 16, type $(1,1,1,1)$ with generators Θ^2, Φ, Ψ, X.

Let an equation $f(\xi)=0$ of degree 32 be irreducible in a field F and have the roots

$$\chi^{a_1}\{\psi^{a_2}\phi^{a_3}\theta^{a_4}\} \qquad (a_1 < 2,\ a_2 < 2,\ a_3 < 2,\ a_4 < 4)$$

satisfying the following relations:

$$\theta^4(i) = \phi^2(i) = \psi^2(i) = \chi^2(i) = i,$$

$$\theta\{\phi\} = \phi\{\theta\}, \qquad \theta\{\chi\} = \chi\{\phi\theta^3\},$$

$$\theta\{\psi\} = \psi\{\phi\theta\}, \qquad \phi\{\chi\} = \chi\{\phi\},$$

$$\phi\{\psi\} = \psi\{\phi\}, \qquad \psi\{\chi\} = \chi\{\psi\}.$$

Then the group constants of the general discussion have the values

$$q = 4 \qquad x = 1,$$

$$p = 2 \qquad z_1 = 1,\ w_1 = 1,$$

$$s = 2 \qquad z_2 = 1,\ w_2 = 0,$$

$$h = 2 \qquad u_1 = 0,\ v_1 = 1,\ y_1 = 3,$$

$$e = a = b = c = d = l = 0, \qquad u_2 = 0,\ v_2 = 1,\ y_2 = 0,$$

$$u_3 = 1,\ v_3 = 0,\ y_3 = 0.$$

The constants c_1 to c_{14} of Dickson's paper are all zero except

$$c_3 = 2, \qquad c_4 = 1, \qquad c_{11} = 1.$$

The table (32) becomes

17. Dickson, "Construction of Division Algebras," Theorem 6.

18. This group was suggested by Dr. J. K. Senior as one which probably cannot be represented with fewer than four generators.

Superscript	ϵ_a	y_a	v_a	u_a	ϵ_b	y_b	v_b	u_b	$\mathcal{F}$	$\mathcal{E}$	$\mathcal{H}$
(1)	ϵ_1	3	1	0	ϵ_2	0	1	0	2	3	$\epsilon_1\epsilon_2\{\phi\theta^3\}$
(2)	ϵ_2	0	1	0	ϵ_3	0	0	1	1	0	$\epsilon_2\epsilon_3\{\phi\}$
(3)	ϵ_1	3	1	0	$\epsilon_2\epsilon_3(\phi)$	0	1	1	2	3	$\epsilon_1\epsilon_2\epsilon_3\{\theta^3\}$
(4)	ϵ_2	0	1	0	ϵ_3	0	0	1	1	0	$\epsilon_2\epsilon_3\{\phi\}$
(5)	1	0	0	0	$\epsilon_2\epsilon_3(\phi)$	0	1	1	1	0	$\epsilon_2\epsilon_3\{\phi\}$
(6)	1	0	0	0	1	0	0	0	0	0	1
(7) 00	1	0	0	0	1	0	0	0	0	0	1
(7) 10	ϵ_2	0	1	0	1	0	0	0	1	0	ϵ_2
(7) 01	1	0	0	0	ϵ_3	0	0	1	0	0	ϵ_3
(8) 000	1	0	0	0	1	0	0	0	0	0	1
(8) 310	Δ	9	3	0	ϵ_2	0	1	0	4	9	$\Delta\epsilon_2\{\phi\theta\}$
(8) 010	1	0	0	0	ϵ_2	0	1	0	1	0	ϵ_2
(8) 001	1	0	0	0	ϵ_3	0	0	1	0	0	ϵ_3
(9)	$\Delta\epsilon_2\{\phi\theta\}$	9	4	0	ρ	0	0	0	4	9	$\Delta\epsilon_2\{\phi\theta\}\rho\{\theta\}$
(10)	ϵ_2	0	1	0	ρ	0	0	0	1	0	$\epsilon_2\rho\{\phi\}$
(11)	ϵ_3	0	0	1	ρ	0	0	0	0	0	$\epsilon_3\rho\{\psi\}$

where $\Delta \equiv \epsilon_1\epsilon_1\{\phi\theta^3\}[\alpha(i)\alpha(\theta^3)]^2\alpha\{\theta\}\alpha\{\phi\}\alpha\{\theta^2\}\alpha\{\phi\theta^2\}\alpha\{\phi\theta^3\}\epsilon_1\{\theta^2\}$.

After evaluating the exponents, we find the conditions in this paper for the present example. They are

(42) $$\mathcal{H}_{21}(i)j^3k^2o^0 = \alpha(\chi)\mathcal{H}^{(1)}(i)j^3k^2o^0,$$

(43) $$\mathcal{H}_{31}(i)j^3k^4o^1 = \epsilon_1(\chi)\mathcal{H}^{(3)}(i)j^3k^2o^1,$$

(44) $$\mathcal{H}_{32}(i)j^0k^1o^1 = \epsilon_2(\chi)\mathcal{H}^{(5)}(i)j^0k^1o^1,$$

(45) $$\mathcal{G}_1{}^{(4)}(i)j^{12}k^4o^0 = g(\chi),$$

(46) $$\mathcal{G}_2{}^{(2)}(i)j^0k^2o^0 = \beta(\chi)\mathcal{G}_1{}^{(0)}(i)j^0k^0o^0,$$

(47) $$\mathcal{G}_3{}^{(2)}(i)j^0k^0o^2 = \sigma(\chi)\mathcal{H}^{(6)}(i)j^0k^0o^0,$$

(50) $$\rho(i)j^0k^0o^0 = \rho(\chi)\mathcal{H}^{(8)}(i)j^0k^0o^0,$$

(59) $$\epsilon_1(\chi)\mathcal{H}^{(0)}j^0k^4o^0 = \rho\mathcal{A}_{01}B_{0,1}j^1k^0o^0,$$

(60) $$\epsilon_2(\chi)\mathcal{H}^{(10)}j^0k^1o^0 = \rho\mathcal{D}_{01}B_{0,0}j^0k^1o^0,$$

(61) $$\epsilon_3(\chi)\mathcal{H}^{(11)}j^0k^0o^1 = \rho(i)j^0k^0o^1.$$

We may now determine completely the associativity conditions for our algebra.

Associativity conditions.

[6] $$g(i) = g(\theta).$$

[14] $$\alpha(i)\alpha(\theta)\alpha(\theta^2)\alpha(\theta^3)g(i) = g(\phi).$$

[15] $$\beta(\phi)=\beta(i).$$

[18] $$\alpha(\phi)\alpha(i)\beta(\theta)=\beta(i).$$

[45] $$\delta_1\delta_1(\phi\theta)\delta_1(\theta^2)\delta_1(\phi\theta^3)\alpha(\theta)\alpha(\theta^2)[\alpha(\theta^3)]^2\alpha(\phi\theta^2)\alpha(\phi\theta^3)[\beta(i)]^2 g(i)=g(\psi).$$

[49] $$\alpha(\psi)\delta_1\delta_2(\phi\theta)=\delta_2\delta_1(\phi)\alpha(i).$$

[54] $$\delta_2\delta_2(\phi)\beta(i)=\beta(\psi).$$

[61] $$\sigma(\psi)=\sigma(i).$$

[69] $$\delta_1\delta_1(\psi)\delta_2(\phi\theta)\sigma(\theta)\beta(\theta)=\sigma(i).$$

[72] $$\delta_2\delta_2(\psi)\sigma(\phi)=\sigma(i).$$

(39) $$\epsilon_2\epsilon_1(\phi)\alpha\alpha(\theta)\alpha(\theta^2)=\alpha(\chi)\epsilon_1\epsilon_2(\phi\theta^3).$$

(43) $$\epsilon_3\epsilon_1(\psi)\delta_1\delta_1(\phi\theta)\delta_1(\theta^2)\delta_2(\phi\theta^3)\alpha(\theta)\alpha(\phi\theta^2)\alpha(\theta^2)\beta(\theta^3)=\epsilon_1(\chi)\epsilon_1\epsilon_2\epsilon_3(\theta^3).$$

(45) $$\epsilon_3\epsilon_2(\psi)\delta_2=\epsilon_2(\chi)\epsilon_2\epsilon_3(\phi).$$

(50) $$\epsilon_1\epsilon_1(\theta^2)\epsilon_1(\phi\theta)\epsilon_1(\phi\theta^3)\alpha\alpha(\phi)\alpha(\phi\theta)$$
$$\times[\alpha(\theta)]^3[\alpha(\theta^2]^4[\alpha(\theta^3)]^4[\alpha(\phi\theta^3)]^2[\alpha(\phi\theta^2)]^2\beta^2 g^3=g(\chi).$$

(54) $$\epsilon_2\epsilon_2(\phi)\beta(i)=\beta(\chi).$$

(58) $$\epsilon_3\epsilon_3(\psi)\sigma(i)=\sigma(\chi).$$

(65) $$\rho(i)=\rho(\chi).$$

(76) $$\epsilon_1(\chi)\epsilon_1\epsilon_1(\phi\theta^3)\epsilon_1(\theta^2)\epsilon_2(\phi\theta)\alpha(\theta)\alpha(\phi\theta^2)$$
$$\times\alpha(\phi\theta^3)\alpha(\phi)\alpha(\theta^2)[\alpha(i)\alpha(\theta^3)\beta(\theta)]^2\rho(\theta)g^2=\rho(i).$$

(77) $$\epsilon_2\epsilon_2(\chi)\rho(\phi)=\rho(i).$$

(81) $$\epsilon_3\epsilon_3(\chi)\rho(\psi)=\rho(i).$$

If we take $g=\beta=\sigma=\rho=\alpha=\delta_1=\delta_2=\epsilon_1=\epsilon_2=\epsilon_3=1$, the algebra over F determined by our equation, $f(\xi)=0$, has the basis

$$i^m j^n k^t o^f E^r \qquad (m\leqq 31,\ n\leqq 3,\ t\leqq 1,\ f\leqq 1,\ r\leqq 1),$$

with $f(i)=0$; and the multiplication table

$$\begin{aligned}
&j^4=1, && jP(i)=P(\theta)j,\quad j^2P(i)=P(\theta^2)j^2,\quad j^3P(i)=P(\theta^3)j^3,\\
&k^2=1, && kP(i)=P(\phi)k,\\
& && kj=jk,\quad kj^2=j^2k,\quad kj^3=j^3k,\\
&o^2=1, && oP(i)=P(\psi)o,\\
& && oj=jko,\quad oj^2=j^2o,\quad oj^3=j^3ko,\\
& && ok=ko,\\
&E^2=1, && EP(i)=P(\chi)E,\\
& && Ej=j^3kE,\quad Ej^2=j^2E,\quad Ej^3=jkE,\\
& && Ek=kE,\\
& && Eo=oE.
\end{aligned}$$

This algebra is associative.

LARGE POSITIVE INTEGERS ARE SUMS OF FOUR OR FIVE VALUES OF A QUADRATIC FUNCTION.*

By GORDON PALL.

1. Professor L. E. Dickson † has, in several papers, considered the representation of all positive integers as sums of $s(\geqq 2)$ values of the function

$$(1) \qquad mx^2/2 + nx/2 + c,$$

$$(m, n, c \text{ integers}; \; m + n \text{ even}; \; m > 0).$$

His formulae give actual limits to p in the following theorem of E. Maillet's.‡

Let $m > 0$, and either: $1°$, mn odd and $(m, n) = 1$; or $2°$, m and n even, $(m + n)/2$ odd, and $(m/2, n/2) = 1$. Then, in case $1°$ every integer p beyond a certain limit [which he does not give explicitly] *depending only on m, n, and c is a sum of four numbers* (1).§ *In case $2°$ the same is true of every sufficiently large odd p.*

Since we can replace p by $p - 4c$ we can suppose that $c = 0$. Without further reference in the sequel we shall suppose $m > 0$ and either $1°$ or $2°$. All cases are seen to reduce to these.

Write

$$(2) \qquad f(x) = mx^2/2 + nx/2.$$

We discuss for $s = 4$ the solvability of the equation

$$(3_s) \qquad p = f(x_1) + \cdots + f(x_s)$$

in integers x_i; and also in integers $x_i \geqq -k$, $(k \geqq 0)$. If better results exist for $s = 5$ we consider (3_5). The most general theorem obtained is

THEOREM 1. *Except when, in case $2°$, $m/2$ is odd and $n/2$ even,* (3_4)

---

* National Research Fellow.

† There are five published ones:

I. *American Journal of Mathematics*, Vol. 50 (1928), pp. 1-48.
II. *Bulletin of the American Mathematical Society*, Vol. 33 (1927), pp. 713-720.
III. *Ibid.*, Vol. 34 (1928), pp. 63-72.
IV. *Ibid.*, Vol. 34 (1928), pp. 205-217.
V. *Proceedings of the American Philosophical Society*, Vol. 66 (1927), pp. 281-286.

‡ *Bulletin de la Société Mathématique de France*, Vol. 23 (1895), pp. 40-49.

§ Maillet does not state condition (12_2) below in its true rôle, as leading to integers $x \geqq 0$. He uses "entier" throughout his paper as though he meant "entier $\geqq 0$."

is solvable in integers $x_i \geqq 0$ *for every* p *exceeding a certain function of* m *and* n. *If* $m/2$ *is odd and* $\geqq 5$, *and* $n/2$ *even, then* (3_4) *is not solvable in integers* x_i *for* 1) *only a finite number of odd* $p > 0$, *and* 2) *an infinite number of even* $p \geqq 0$. *The values* p *in* 2) *are such that* $2mp + n^2$ *forms a finite number, say* w, *of progressions*

$$(4) \qquad 4^{lh}t_v (v = 1, \cdots, w), \quad h = 0, 1, 2, \cdots,$$

where the t_v *are even integers and* l *is the least positive integer such that* $2^l \equiv \pm 1 \pmod{m/2}$.

The equation (3_s) is equivalent to

$$(5_s) \qquad 8mp + sn^2 = (2mx + n)^2.$$

Hence our theorems are equivalent to facts about sums of squares of numbers in an arithmetic progression.

For example, for $m/2 = 5$, consider

$$(6_j) \qquad 5p + 4j^2 = (5x_1 + j)^2 + \cdots + (5x_4 + j)^2,$$

where $j = 1$ or 2. We shall see that the only integers $5p + 4 \geqq 0$ for which (6_1) is not solvable are 9, 29, 59, 79, and $4^{2h}t$, where

$$(7) \qquad t = 14, 24, 44, 94, 184, \; h \geqq 0.$$

And the only integers $5p + 16 \geqq 0$ for which (6_2) is not solvable in integers x_i are 1, 11, 41, 51, 101, and $4^{2h}t$,

$$(8) \qquad t = 6, 46, 56, 176, 376, \; h \geqq 0.$$

Let $j = 1$ or 2. We shall in Section 4 find all integers $q \geqq 0$ such that

$$(9_j) \qquad 3q + 4 = (3x_1 + j)^2 + \cdots + (3x_4 + j)^2$$

is not solvable in integers $x_i \geqq -k$; $(k \geqq 0)$. It is remarkable that, while there are for any $k \geqq 0$, infinitely many positive $(3q + 4)$'s such that (9_j) is not solvable in integers $x_i \geqq -k$, yet every $3q + 4 \geqq 4$ is a sum of four squares prime to 3.

It is an obvious corollary of Theorem 1 that (3_5) is solvable in integers x_i for every sufficiently large integer p. This is seen to be true in integers $x_i \geqq 0$. We tabulate in Sections 3 and 4 a number of the simplest cases (5_3), (5_4), and (5_5).

2. We shall use two extensions of the Cauchy lemma on the solvability in integers x_i of the equations

(10) $$a = x_1^2 + \cdots + x_4^2, \quad b = x_1 + \cdots + x_4.$$

Proofs are given in the writer's paper, "Simultaneous Quadratic and Linear Representation", to appear shortly in the *Quarterly Journal of Mathematics*.

Let Λ denote the *form* $4^h(8v+7)$, where h and v are integers, $h \geqq 0$.

LEMMA 1. *Integral solutions x_i of (10) exist if and only if*

(11) $$a \equiv b \,(mod\, 2), \quad 4a - b^2 \neq \Lambda, \quad 4a \geqq b^2.$$

None of the x_i in (10) can be negative if

(12) $$b \geqq 0, \quad b^2 + 2b + 4 > 3a.$$

If a, b are even, solutions in integers $x_i \geqq 0$ exist if (11_2), (11_3), (12), and merely

(13) $$3b^2 + 8b + 16 > 8a.$$

LEMMA 2. *If $s = 5$, 6, or 7, integers x_w satisfying*

(14) $$a = x_1^2 + \cdots + x_s^2, \quad b = x_1 + \cdots + x_s$$

exist if and only if

(15) $$a \equiv b \,(mod\, 2), \quad sa \geqq b^2.$$

Such integers $\geqq 0$ exist if, in addition to (15),

(16) $$b \geqq 0, \quad b^2 \geqq 3a - 5.^*$$

3. Since $(m, n) = 1$ or 2 every integer p is of the form

(17) $$p = (am + bn)/2,$$

where a, b are integers. With Dickson we write $t = (m+n)/2$. Then (17) may be written as

(18) $$p = m(a-b)/2 + tb.$$

LEMMA 3. *Let r denote one of the numbers $0, 1, \cdots, m-1$. Then*

(19) $$p \equiv rt \,(mod\, m)$$

and $b \equiv r \,(mod\, m)$ imply that (17) defines an integer a of the same parity as b. Conversely, if $b \equiv r \,(mod\, m)$ and $a \equiv b \,(mod\, 2)$, (17) implies (19).

3.1. Taking the simplest case first we find all integers p permitting integer solutions a, b of

---

* If a, b are odd, (16_2) may be replaced by (12_2). If a, b are even, (16_2) may be replaced by (11_2) and (13).

$$(20)\qquad p=(am+bn)/2,\quad sa\geqq b^2,\quad a\equiv b\pmod 2),$$

where s is given and $\geqq 1$. Conditions (20_1) and (20_2) are given equivalent to (20_1) and

$$(21)\qquad p\geqq g(b)/2,\qquad g(b)=nb+mb^2/s.$$

Let b_r denote the integer $b\equiv r\pmod m$ for which $g(b)$ is least. (By calculus, $2m\mid b_r\mid\leqq m^2+s\mid n\mid$.) Then the integers required are all p such that

$$(22)\qquad p\geqq g(b_r)/2,\quad p\equiv rt\pmod m,\quad (r=0,1,\cdots,m-1).$$

By lemma 2 we have

LEMMA 4. *If $s=5$, 6, or 7, the values p in (22) form the class of all integers p for which (3_s), or (5_s), is solvable in integers x_i.*

We list all cases in which the number of integers $8mp+5n^2\geqq 0$ for which (5_5) has no solution in integers x_i is $\leqq 5$, and in which there is not a better result (given later) for $s=4$.

	Solvable for every $p\geqq 0$ except the following:
$56p+45=\sum_5(14x+3)^2$,	$p=1,3$;
$56p+125=\sum_5(14x+5)^2$,	none ;
$5p+20=\sum_5(5x+2)^2$,	$p=6,7,8,17$;
$16p+45=\sum_5(8x+3)^2$,	$p=6$;
$16p+5=\sum_5(8x+1)^2$,	$p=1,2,4,7$;
$72p+125=\sum_5(18x+5)^2$,	$p=1,3,5,12$;
$72p+245=\sum_5(18x+7)^2$,	$p=6,7$.

3.2. Second we wish all integers p for which integers a, b exist satisfying

$$(23)\qquad (17),\ sa\geqq b^2\geqq \tau a-\beta b-\gamma,\quad a\equiv b\pmod 2),\quad b\geqq -k,$$

where τ, β, γ, s, k are given rational numbers, $0<\tau<s$.

Conditions (23_1) and (23_2) are equivalent to (23_1) and

$$(24)\qquad I_b:\ g(b)/2\leqq p\leqq h(b)/2,$$

$$g(b)=nb+mb^2/s,\quad h(b)=nb+m(b^2+\beta b+\gamma)/\tau.$$

The intervals I_b for $b\geqq -k$ and $\equiv r\pmod m$ give all values $p\equiv rt\pmod m$ for which (23) are solvable.

3.21. For simplicity we now suppose $n>-m$. Then $g(b)$ and $h(b)$

are monotone increasing as soon as $b \geqq \sigma/2$, $\sigma = \max(s-1, \tau-\beta-1)$. For any integer $d \geqq 1$ let $y = y(d)$ be the least integer $\geqq \sigma/2$ such that

$$(25) \qquad g(z+d) \leqq h(z) \text{ for every integer } z \geqq y.$$

Then, clearly, every p such that

$$(26) \qquad p \geqq g(y+d-1)/2$$

belongs to at least d consecutive intervals I_b.

On completing the square for z, (25) becomes

$$(27) \qquad [(s-\tau)y-(\tau d-\beta s/2)]^2 \geqq Z, \quad (s-\tau)(y+1/2) \geqq \tau d-\beta s/2,$$

where $\quad Z = s\tau(d^2-\beta d+\gamma)-s^2(\gamma-\beta^2/4)+s\tau(s-\tau)dn/m.$

3.22. If $s=5$, $k=0$, $\beta=0$, $\gamma=5$, $\tau=3$, then (23) are only sufficient for the solvability of (5_5) in integers $x_i \geqq 0$. It is clear that (5_5) is solvable in integers $x_i \geqq 0$ for every p satisfying (26) with $d=m$. The values p below this limit must be considered separately. This is practicable only for a small m; we can make use also of the intervals I_b for $b < y(m) + m - 1$. Also, if b is odd, we can use (12_2) in place of (16_2).

3.3. Third let $s=4$ and adjoin the condition (11_2) to (20). If b_r is odd condition (11_2) is superfluous, since, for odd a and b, $4a-b^2 \neq \Lambda$.

LEMMA 5. *For a certain r let b_r be even, and let b' denote that one of $b_r \pm m$ for which $g(b')$ is least. If m is odd, or if $m/2$ is even, or if $m/2$ is odd and p such that*

$$(28) \qquad mp/2+n^2/4 \not\equiv 0 \pmod 8,$$

then (5_4) is solvable in integers x_i for $p \equiv rt \pmod m$ and $\geqq g(b')/2$.

Remark. The only values $p \equiv rt$ and $< g(b')/2$ for which (5_4) can possibly be solvable are in the interval

$$(29) \qquad g(b_r)/2 \leqq p < g(b')/2.$$

Since (20) is solvable for these values p only with $b=b_r$ we need merely exclude among these numbers p all for which

$$4(2p-b_r n)/m-(b_r)^2 = \Lambda$$

to obtain all values $p \equiv rt$ and $< g(b')/2$ for which (5_4) is solvable in integers.

If m is odd, b' is odd, and lemma 5 obvious.

If $m/2$ is even, $n/2$ is odd, and one of $a_1 = (2p - nb_r)/m$, $a_2 = (2p - nb')/m$ is $\equiv 2 \pmod 4$. But $a \equiv 2 \pmod 4$, b even, imply $4a - b^2 \neq \Lambda$.

If $m/2$ is odd, $n/2$ is even, and we can choose k uniquely modulo 2 so that

$$m(b_r + km) + 2n = 4Q,\ Q \text{ odd}.$$

Hence, by (17) and (28),

$$m^2(4a - b^2) = 8mp + 4n^2 - (mb + 2n)^2 = 16(8v + 1, 3, \text{or } 5),$$

for one of $b = b_r$, b'. Lemma 5 follows.

In this third case we treat together * the system of equations

$$(30) \qquad mp/2 + n_j^2/4 = \textstyle\sum_i (mx_i/2 + n_j/4)^2 \qquad (j = 1, \cdots, l),$$

where l is the least integer > 0 such that $2^l \equiv \pm 1 \pmod{m/2}$, and the $n_j/2$ form, apart from sign, a representative set of even residues modulo $m/2$ of $2^h \cdot n/2$ $(h = 0, 1, 2, \cdots)$. For any $h \geqq 0$ and j we can choose a unique g from $(1, \cdots, l)$ so that

$$(31) \qquad 2^h n_g \equiv \pm n_j \pmod m.$$

If, for some h and j, $4^{-h}(mp/2 + n_j^2/4)$ is an even integer, every square in the representations of $mp/2 + n_j^2/4$ as a sum of four squares is divisible by 4^h. Then, by (31),

$$(32) \qquad mp/2 + n_j^2/4 = 4^h(mq/2 + n_g^2/4),$$

where q is an even integer; and for any x_i of (30_j),

$$(33) \qquad mx_i/2 + n_j/4 = 2^h(my_i/2 + n_g/4),$$

where y_i is an integer.

All even values p for which any one of the l equations (30) is solvable in integers x_i can therefore be deduced from a knowledge of all values p satisfying

$$(34) \qquad mp/2 + n_j^2/4 \not\equiv 0 \pmod 8,\ p \geqq -(n_j)^2/2m,$$

for which each equation is solvable.

It is clear, also, that if some equation (30_g) is not solvable in integers x_i for one even value $p = q$ satisfying (34), then each equation (30_j) has

* Instead we can consider just the single equation (5_4) for all even values p such that $mp/2 + n^2 \not\equiv 0 \pmod{2.4^l}$.

no solution in integers x_i for infinitely many values $p \geqq 0$, indeed for all p defined by (32) with h as in (31).

We prove for $m/2 \geqq 5$, $m/2$ odd, that (5_4) has no solution in integers x_i for some even $p \geqq -n^2/2m$. We can suppose that $0 < n < m$. Hence the three least squares $(mx_i/2 + n/4)^2$ are the squares of

$$x = n/4,\ y = m/2 - n/4,\ z = m/2 + n/4;\ x^2 < y^2 < z^2.$$

Their even sums by four possibly $\leqq z^2 + 3x^2$ are

$$4x^2,\ 2y^2 + 2x^2,\ 4y^2,$$

three in number. But at least four even values p satisfy

$$0 \leqq pm/2 + n^2/4 \leqq z^2 + 3x^2.$$

The results for (6) may be verified in about ten minutes by means of (22) with $m = 10$, $n = 4$ or 8, whence $t = 7$ or 9; and $b_r = 1, 3, -5, -3, -1$, $(r = 1, 3, 5, 7, 9)$; and by means of lemma 5 and succeeding remark with $(b_r, b') = (0, -10)$, $(2, -8)$, $(4, -6)$ or $(-6, 4)$, $(-4, 6)$, $(-2, 8)$ $(r = 0, 2, 4, 6, 8)$.

	The only integers $p \geqq 0$ for which equation is not solvable in integers x are
$40p + 4 = \sum_4(10x+1)^2$,	$p = 1$;
$40p + 36 = \sum_4(10x+3)^2$,	none ;
$56p + 4 = \sum_4(14x+1)^2$,	$p = 1, 2, 5$;
$56p + 36 = \sum_4(14x+3)^2$,	$p = 1, 3, 25$;
$56p + 100 = \sum_4(14x+5)^2$,	$p = 5, 23$;
$16p + 4 = \sum_4(8x+1)^2$,	$p = 1, 2, 4, 7, 30$;
$16p + 36 = \sum_4(8x+3)^2$,	$p = 5, 6, 26$.

3.4. Fourth we adjoin (11_2) to (23) for $s = 4$. From the intervals I_b must be excluded all values p such that

$$4(2p - bn)/m - b^2 = \Delta.$$

As in 3.3, we see that, if m is odd, or $m/2$ even, or if $m/2$ is odd and p satisfies (28), then (5_4) is solvable in integers $x_i \geqq 0$ if (26) holds with $d = 2m$, $\beta = 2$, $\gamma = \tau = 3$.

4. For $s = 3$ we note that each of the equations

$$8p + 3 = \sum_3(4x+1)^2, \quad 24p + 3 = \sum_3(6x+1)^2,$$

is solvable in three integers x for every $p \geqq 0$. Not as trivial is the fact that

every positive multiple $3N$ of 3 is a sum of three squares prime to 3, if $3N \neq \Lambda$. This is evident unless $3N$ is also a multiple of 9. Then it follows from the identity

$$9(x^2+y^2+z^2) = (2x+2y-z)^2 + (2z+2x-y)^2 + (2y+2z-x)^2.$$

For any $p \geqq 0$ we can write $3p+4 = 4^w(3q+4)$, where $q \not\equiv 0 \pmod 4$ and $q \geqq 0$, or $q = 0$. Write $3N = 3q$ unless $3N = \Lambda$, then $3N = 3q+3$. Add 4 or 1 respectively. Hence: *every* $3p+4$ *is a sum of four squares prime to* 3, one of the squares being 1 or 4. It is of interest to recall Pepin's formula for the number of representations in $x^2+9y^2+9z^2+9t^2$. A corollary is that if p is even, the number of representations of $3p+4$ as a sum of four squares prime to 3 is equal to $16\zeta'_1(3p+4)$, sixteen times the sum of the odd divisors of $3p+4$.

To discuss (9_j) for integers $x_i \geqq 0$ we use the formulae of Section 3 with $m=6$, $n=4j$ ($j=1$ or 2), $s=4$, $q=p+4(j-1)$. If $d=6$, (27_1) becomes

$$(35)\qquad (y-14)^2 \geqq 292+48j; \text{ or } (y-8)^2 \geqq 121+32j;$$

according as

$$(36)\qquad \tau=\gamma=3,\ \beta=2, \text{ or } \tau=\beta=2\tfrac{2}{3},\ \gamma=5.$$

Hence $y = 32+j$, $20+j$ respectively, and, by (26), every $p \geqq \mathbf{g}(y+5)/2$ belongs to at least six intervals I_b. If p is odd or double of an odd this means the existence of a suitable a which is odd or double of an odd. Hence (9_j) is solvable in integers $x_i \geqq 0$ for every p such that

$$(37)\qquad p \geqq 1021+138j,\ p \text{ odd};\quad p \geqq 466+96j,\ p \equiv 2 \pmod 4).$$

Setting $m=6$, $t=3+2j$ in Dickson's table I * we get a list T_j of all sums $\leqq 1335+28j$ of four values of $f(x)$ for integers $x \geqq 0$. This was checked independently to $489+44j$ by setting $m=6$, $t=1$ in Dickson's table III.† From T_1 and T_2 we get

LEMMA 6. *The following are all odd* $q>0$ *such that*

$$(38)\qquad 3q+4 \neq \Sigma_4(1^2, 4^2, 7^2, 10^2, \cdots):$$

(39) $q = 1, 3, 7, 9, 11, 13, 17, 19, 23, 25, 27, 29, 35, 39, 41, 45, 47, 51, 55, 57, 63, 67, 69, 73, 75, 79, 83, 91, 97, 103, 107, 109, 113, 119,$

* *L. c.* II, pp. 718, 719.

† *L. c.* IV, pp. 207, 208.

129, 131, 137, 143, 147, 149, 159, 165, 183, 189, 195, 211, 235, 239, 247, 263, 275, 305, 321, 339, 345, 403, 509, 585, 643.

The following are all odd $q > 0$ such that

(40) $$3q + 4 \neq \sum_4(2^2, 5^2, 8^2, 11^2, \cdots):$$

(41) $q = 1, 3, 5, 7, 9, 13, 15, 17, 19, 21, 23, 27, 29, 33, 35, 37, 39, 41, 47, 49, 53, 55, 59, 61, 65, 67, 69, 73, 79, 81, 85, 87, 91, 93, 97, 101, 105, 111, 117, 123, 125, 129, 131, 135, 137, 149, 151, 155, 161, 167, 169, 173, 181, 187, 205, 211, 217, 223, 229, 237, 261, 269, 273, 279, 285, 293, 305, 335, 341, 355, 373, 379, 409, 441, 551, 631, 693, 1169.$

LEMMA 7. *All positive integers $q \equiv 2 \pmod 4$ satisfying (38) are*

(42) $q = 2, 6, 14, 18, 22, 30, 34, 46, 50, 58, 62, 74, 78, 86, 102, 114, 126, 142, 154, 206, 258, 270, 334, 398.$

All such satisfying (40) are

(43) $q = 2, 6, 10, 14, 22, 26, 30, 34, 42, 46, 54, 62, 66, 74, 78, 86, 94, 98, 110, 118, 130, 142, 154, 162, 174, 198, 230, 254, 286, 298, 366, 434.$

For $q \equiv 0 \pmod 8$ we use $d = 12$ which is effective, since an integer $a \equiv 4 \pmod 8$ will appear from $p = 3a + 2jb$ for one of $b \equiv 2, 6, 10 \pmod{12}$. Under (36_2), (27_1) yields

$$(y_{12} - 20)^2 \geqq 673 + 64j, \quad y_{12} = 47 + j.$$

Hence (9_j) is solvable in integers $x_i \geqq 0$ for all $p \equiv 4 - 4j \pmod 8$ and $\geqq 2524 + 208j$. Now I_{58}, I_{46}, I_{34} are respectively

$$2523 \leqq p - 116j \leqq 3964, \quad 1587 \leqq p - 92j \leqq 2524, \quad 867 \leqq p - 68j \leqq 1408;$$
I_{54}, I_{42} are $2187 \leqq p - 108j \leqq 3448, \quad 1323 \leqq p - 84j \leqq 2116;$
I_{50}, I_{38} are $1875 \leqq p - 100j \leqq 2968, \quad 1083 \leqq p - 76j \leqq 1744.$

Above the limit $1335 + 28j$ of T_j and outside the intervals I_b just given we find that there remain just the following values

$p \equiv b(3 + 2j) \pmod 6$ and $\equiv 4(1 - j) \pmod 8$:

$$\begin{array}{ll} p = 1428 + 68j + 24v & (v = 0, 1, \cdots, 6 + j); \\ \quad 2532 + 92j + 24v & (v = 0, j - 1); \\ \quad 2124 + 84j + 24v & (v = 0, 1, \cdots, 2 + j); \\ \quad 1332 + 36j + 24v & (v = 0, \cdots, 2j - 1); \\ \quad 1764 + 76j + 24v & (v = 0, 1, \cdots, 4 + j). \end{array}$$

All these numbers were located in the extension of T_j obtained by setting $m = 6$, $t = 3 + 2j$ in Dickson's table on pp. 13, 14 of I, except $p = 2624$ if $j = 1$, and $p = 2716$ and 2740 if $j = 2$. But, $2624 = f(29) + f(3) + 2f(1)$, $2716 = f(27) + f(11) + 2f(1)$, $2740 = f(27) + 2f(7) + f(5)$.

Hence table T_j gives

LEMMA 8. *All integers* $q \geqq 0$, $\equiv 0 \pmod 8$, *and satisfying* (40), *are*

(44) 8, 24, 40, 80, 96, 160, 200, 216, 424, 520, 1224.

All integers $q \geqq 0$, $\equiv 0 \pmod 8$, *and satisfying* (40), *are*

(45) 0, 8, 16, 40, 48, 56, 72, 80, 104, 112, 144, 168, 192, 224, 256, 280, 416, 440.

From lemma 6 we can verify

LEMMA 9. *All odd* $q > 0$ *such that* (9_1) *is not solvable in* $x_i \geqq -1$ *are*

(46) 9, 13, 25, 29, 41, 45, 47, 69, 75, 79, 97, 109, 149, 165, 189, 235, 305, 509.

For each of these (9_1) *is solvable in integers* $x_i \geqq -2$. *All odd* $q > 0$ *such that* (9_2) *is not solvable in integers* $x_i \geqq -1$ *are*

(47) 5, 7, 13, 15, 19, 27, 35, 39, 47, 53, 55, 79, 85, 91, 93, 111, 123, 167, 187, 211, 223, 261, 279, 285, 335, 551; 33, 59, 129.

For the numbers preceding the semicolon (9_2) *is solvable in integers* $x_i \geqq -2$. *For* $q = 33, 59, 129$ *it is first solvable in integers* $x_i \geqq -3$.

Consider the class of the sixty-nine numbers $t = 3q + 4$, where q runs over the even integers q of (42), (43), (44), and (45). This is the set of all even numbers $\equiv 1 \pmod 3$ and $\not\equiv 0 \pmod 8$, such that the list of all representations

$$t = 3q + 4 = y_1^2 + y_2^2 + y_3^2 + y_4^2, \qquad \text{all } y_i \text{ prime to } 3, \quad y_1 \geqq y_2 \geqq y_3 \geqq y_4 > 0, \tag{48}$$

does not contain two representations (y_i) and (z_i) such that each $y_i \equiv 1$ and each $z_i \equiv 2 \pmod 3$.

If $k \geqq 0$ and q even, the equation

$$4^k(3q + 4) = (3x_1 + j)^2 + \cdots + (3x_4 + j)^2 \tag{49}$$

will not be solvable in integers $x_i \geqq -k$ if and only if in every representation (48),

(50) some y_i satisfies $2^k y_i > 3k - j$ and $2^k y_i \not\equiv j \pmod 3$.

For else we can choose x_i from $3x_i + j = \pm 2^h y_i$.

An application of (50) to the list of all representations (48) for each t gives all values h for which (49) is not solvable in integers $x_i \geqq -k$. For example, for

(51) $t = 4, 34, 52, 130, 148, 172, 202, 286, 298, 316, 340, 358, 394, 436, 490, 526, 580, 598, 676, 694, 766, 772, 844, 862, 898, 1102, 1252, 1306,$

the values $q = (t-4)/3$ appear in (43) or (45) and not in (42) or (44). Hence each t possesses a representation (48) with each $y_i \equiv 1$ and none with each $y_i \equiv 2 \pmod 3$. To exclude this representation, (50) requires $2^h \not\equiv j \pmod 3$, that is, $h \equiv j \pmod 2$. Also, (51) possess respectively the partitions (y_1, y_2, y_3, y_4) which follow:

$(1,1,1,1)$, $(5,2,2,1)$, $(5,5,1,1)$, $(11,2,2,1)$, $(11,5,1,1)$, $(11,5,5,1)$ $(14,2,1,1)$, $(14,8,5,1)$, $(17,2,2,1)$, $(17,5,1,1)$, $(17,5,5,1)$, $(17,8,2,1)$, $(14,14,1,1)$, $(17,11,5,1)$, $(20,8,5,1)$, $(20,11,2,1)$, $(23,5,5,1)$, $(23,8,2,1)$, $(23,11,5,1)$, $(20,17,2,1)$, $(26,8,5,1)$, $(23,11,11,1)$, $(29,1,1,1)$, $(26,11,8,1)$, $(26,14,5,1)$, $(29,14,8,1)$, $(35,5,1,1)$, $(29,20,8,1)$,

in which every $y_i > 1$ is $\equiv 2 \pmod 3$. Since we already have $2 \cdot 2^h \equiv j \pmod 3$, (50) requires finally that

(52) $$2^h > 3k - j, \quad h \equiv j \pmod 2;$$

and it is evident that (52) serves to assure (50) for every further representation (48) of any t in (51).

In some cases it is necessary to consider more than one representation (48) of t, and sometimes all representations. If

(53) $t = 58, 154, 178, 292, 310, 346, 382, 604, 622, 778, 814, 1006, 1198, 1276, 3676,$

then q belongs to (42) or (44), and not to (43) or (45). Also they have the respective partitions

$(7,2,2,1)$, $(10,7,2,1)$, $(13,2,2,1)$, $(16,4,4,2)$, $(16,7,2,1)$, $(13,13,2,2)$, $(19,4,2,1)$, $(22,10,4,2)$, $(19,16,2,1)$, $(25,10,7,2)$, $(25,13,4,2)$, $(28,13,7,2)$, $(28,19,7,2)$, $(34,10,4,2)$, $(52,22,22,2)$,

in which each $y_i \neq 2$ is $\equiv 1 \pmod 3$. The condition on h is seen to be

(54) $$2 \cdot 2^h > 3k - j, \quad h \not\equiv j \pmod 2.$$

If t possesses a representation (48) with one or more $y_i = 1$ and the remaining $y_i \equiv 2$ and also a representation with one or more $y_i = 2$ and the remaining $y_i \equiv 1$ (mod 3), then (50) demands

(55) $\qquad 2^h > 3k - j$, or $2 \cdot 2^h > 3k - j \geqq 2^h$ and $h \not\equiv j \pmod 2$.

If also q belongs to both (42) and (43), or (44) and (45), then (55) assures (50) for all remaining representations. The values t for which these properties hold are

(56) $\qquad t = 10, 28, 70, 124, 190, 226, 262, 430, 466,$

the respective partitions being

(2, 2, 1, 1); (5, 1, 1, 1) = (4, 2, 2, 2); (8, 2, 1, 1) = (7, 4, 2, 1); (11, 1, 1, 1) = (10, 4, 2, 2); (13, 4, 2, 1) = (11, 8, 2, 1); (14, 5, 2, 1) = (13, 7, 2, 2); (16, 2, 1, 1) = (14, 8, 1, 1); (20, 5, 2, 1) = (19, 7, 4, 2); (20, 8, 1, 1) = (19, 10, 2, 1).

All representations (48) of 17 values t:

94 = (8, 5, 2, 1) = (7, 5, 4, 2);
244 = (14, 4, 4, 4) = (13, 7, 5, 1) = (13, 5, 5, 5) = (11, 11, 1, 1) = (11, 7, 7, 5) = (10, 8, 8, 4);
22 = (4, 2, 1, 1);
106 = (10, 2, 1, 1) = (8, 5, 4, 1) = (7, 7, 2, 2) = (7, 5, 4, 4);
238 = (14, 5, 4, 1) = (13, 8, 2, 1) = (13, 7, 4, 2) = (11, 10, 4, 1) = (11, 8, 7, 2) = (10, 8, 7, 5);
46 = (5, 4, 2, 1);
142 = (11, 4, 2, 1) = (10, 5, 4, 1) = (8, 7, 5, 2);
82 = (8, 4, 1, 1) = (7, 5, 2, 2) = (7, 4, 4, 1) = (5, 5, 4, 4);
166 = (11, 5, 4, 2) = (10, 8, 1, 1) = (10, 7, 4, 1) = (10, 5, 5, 4) = (8, 7, 7, 2);
220 = (14, 4, 2, 2) = (13, 7, 1, 1) = (13, 5, 5, 1) = (11, 7, 7, 1) = (11, 7, 5, 5) = (10, 10, 4, 2);
334 = (17, 5, 4, 2) = (16, 7, 5, 2) = (14, 11, 4, 1) = (14, 8, 7, 5) = (13, 10, 8, 1) = (13, 10, 7, 4) = (11, 10, 8, 7);
76 = (8, 2, 2, 2) = (7, 5, 1, 1) = (5, 5, 5, 1);
484 = (20, 8, 4, 2) = (19, 11, 1, 1) = (19, 7, 7, 5) = (17, 13, 5, 1) = (17, 11, 7, 5) = (16, 14, 4, 4) = (16, 10, 8, 8) = (13, 13, 11, 5) = (11, 11, 11, 11);
652 = (25, 5, 1, 1) = (23, 11, 1, 1) = (23, 7, 7, 5) = (22, 10, 8, 2) = (19, 17, 1, 1) = (19, 13, 11, 1) = (19, 11, 11, 7) = (17, 17, 7, 5)

$= (17, 13, 13, 5) = (17, 11, 11, 11) = (16, 14, 14, 2) = (16, 14, 10, 10)$
$= (14, 14, 14, 8)$;

$1564 = (38, 10, 4, 2) = (37, 13, 5, 1) = (37, 11, 7, 5) = (35, 17, 7, 1)$
$= (35, 17, 5, 5) = (35, 13, 13, 1) = (35, 13, 11, 7) = (34, 20, 2, 2)$
$= (34, 14, 14, 4) = (31, 23, 7, 5) = (13, 19, 11, 11) = (31, 17, 17, 5)$
$= (29, 25, 7, 7) = (29, 23, 13, 5) = (29, 19, 19, 1) = (28, 26, 10, 2)$
$= (28, 22, 14, 10) = (26, 26, 14, 4) = (26, 22, 20, 2) = (25, 25, 17, 5)$
$= (25, 23, 19, 7) = (25, 23, 17, 11) = (25, 19, 17, 17)$
$= (22, 22, 20, 14)$;

$508 = (22, 4, 2, 2) = (20, 10, 2, 2) = (19, 11, 5, 1) = (19, 7, 7, 7)$
$= (17, 13, 7, 1) = (17, 13, 5, 5) = (17, 11, 7, 7) = (14, 14, 10, 4)$
$= (13, 13, 13, 1) = (13, 13, 11, 7)$;

$1324 = (35, 7, 7, 1) = (35, 7, 5, 5) = (34, 10, 8, 2) = (32, 14, 10, 2)$
$= (32, 10, 10, 10) = (31, 19, 1, 1) = (31, 17, 7, 5) = (31, 13, 13, 5)$
$= (31, 11, 11, 11) = (29, 19, 11, 1) = (29, 17, 13, 5) = (26, 22, 10, 8)$
$= (26, 16, 14, 14) = (25, 25, 7, 5) = (25, 23, 13, 1) = (25, 23, 11, 7)$
$= (25, 19, 17, 7) = (25, 19, 13, 13) = (25, 17, 17, 11)$
$= (22, 22, 16, 10)$.

Applying (50) and collecting the preceding results we have

THEOREM 2. *Let $j = 1$ or 2, $k \geqq 0$. The odd values q for which (9_j) has no solution in integers $x_i \geqq -k$ are seen in lemmas 6 and 9. The even values q are $(4^h t - 4)/3$ where:*

1) *t is given by* (51); *h by* (52);
2) *t by* (53); *h by* (54);
3) *t by* (56); *h by* (55);
4) $t = 94, 244$; $2^h > 3k - j$, *or* $5 \cdot 2^h > 3k - j \geqq 2^h$ *and* $h \not\equiv j \pmod 2$;
5) $t = 22, 106, 238$; $2 \cdot 2^h > 3k - j$, *or* $4 \cdot 2^h > 3k - j \geqq 2 \cdot 2^h$ *and* $h \equiv j$;
6) $t = 46, 142$; $4 \cdot 2^h > 3k - j$, *or* $5 \cdot 2^h > 3k - j \geqq 4 \cdot 2^h$ *and* $h \not\equiv j$;
7) $t = 82, 166, 220, 334$; $4 \cdot 2^h > 3k - j$, $h \equiv j \pmod 2$;
8) $t = 76, 484, 652, 1564$; $5 \cdot 2^h > 3k - j$, $h \not\equiv j$;
9) $t = 508, 1324$; $7 \cdot 2^h > 3k - j$, $h \equiv j$.

CALIFORNIA INSTITUTE OF TECHNOLOGY.

ON STIELTJES CONTINUED FRACTIONS.

By J. SHOHAT (JACQUES CHOKHATE).

The object of this paper is to investigate the relation of the denominators of the *odd* convergents of the continued fraction

$$\int_a^b \frac{d\psi(y)}{x-y} = \frac{1/}{/l_1 x} + \frac{1/}{/l_2} + \frac{1/}{/l_3 x} + \frac{1/}{/l_4} + \cdots$$

to the corresponding system of orthogonal and normal Tchebycheff polynomials, and to show their application to the theory of mechanical quadratures.

1. Let $\psi(x)$ be a bounded non-decreasing function, defined on the—finite or infinite—interval (a, b), such that

$$\textit{all moments } \alpha_i = \int_a^b x^i d\psi(x) \quad (i = 0, 1, \cdots) \textit{ exist, with } \alpha_0 > 0.$$

We have then a corresponding system of orthogonal and normal Tchebycheff polynomials

$$(1) \quad \begin{gathered} \phi_n(x) \equiv \phi_n(x;\, a, b;\, d\psi) \equiv \phi_n(x;\, d\psi) = a_n x^n + \cdots \quad (n = 0, 1, \cdots) \\ a_n \equiv a_n(d\psi) > 0, \end{gathered}$$

uniquely determined by means of the relations

$$(2) \quad \int_a^b \phi_m(x)\phi_n(x)\, d\psi(x) = 0 \ (m \neq n),\ 1\ (m = n) \quad (m, n = 0, 1, \cdots).$$

We shall use also the orthogonal polynomials

$$(3) \quad \Phi_n(x) \equiv \Phi_n(x;\, a, b;\, d\psi) \equiv \Phi_n(x;\, d\psi) \equiv \phi_n(x)/a_n = x^n + \cdots \quad (n = 0, 1, \cdots),$$

which, by virtue of (1), are uniquely determined by means of

$$(4) \quad \int_a^b \phi_n(x)\, G_{n-1}(x)\, d\psi(x) = 0,$$

where $G_s(x) = \sum_{i=0}^{s} g_i x^i$ stands in general for an arbitrary polynomial of degree $\leq s$ (subject, in some cases, to certain explicitly stated conditions).

The polynomials (3), are, as is known, the denominators of the successive convergents of the "associated" continued fraction *

$$(5)\qquad \int_a^b \frac{d\psi(y)}{x-y} = \frac{\lambda_1/}{/x-c_1} - \frac{\lambda_2/}{/x-c_2} - \cdots \qquad (\lambda_i,\, c_i = \text{const.}),$$

which always exists, $\psi(x)$ being of the above type. (5), in turn, is intimately connected with the "corresponding" continued fraction *

$$(6)\ \int_a^b \frac{d\psi(y)}{x-y} = \frac{1/}{/l_1x} + \frac{1/}{/l_2} + \frac{1/}{/l_3x} + \frac{1/}{/l_4} + \cdots \equiv \frac{b_1/}{/x} - \frac{b_2/}{/1} - \frac{b_3/}{/x} - \frac{b_4/}{/1} - \cdots \dagger$$
$$\left((l_i,\, b_i = \text{const.},\ b_i = -\frac{1}{l_{i-1}l_i}\right),$$

which exists, if $a \geqq 0$. In fact, denoting by $A_n(x)/B_n(x)$ $(n = 0, 1, \cdots)$ the successive convergents to (6), we have (1):

$$(7)\qquad B_{2n+\epsilon}(x) = l_1l_2\cdots l_{2n+\epsilon}x^{n+\epsilon} + \cdots\ (B_{2n+1}(0) = 0) \qquad (\epsilon = 0, 1)$$

$$(8)\qquad \int_a^b B_{2n+\epsilon}(x)x^k d\psi(x) = 0 \qquad (k = 0, 1, \cdots, n-1)$$

$$(9)\qquad \int_a^b B_{2n}^2(x)\,d\psi(x) = \frac{1}{l_{2n+1}}\,;\quad \int_a^b x\left(\frac{B_{2n+1}(x)}{x}\right)^2 d\psi(x) = -\frac{1}{l_{2n+2}}\ .$$

(8), with $\epsilon = 0$, shows, by virtue of (3, 4, 9), that

$$(10)\qquad B_{2n}(x) \equiv \Phi_n(x;\, d\psi)/\Phi_n(0;\, d\psi) \equiv \phi_n(x;\, d\psi)/\phi_n(0;\, d\psi),$$
for
$$(11)\qquad l_{2n+1} = \Phi_n^2(0;\, d\psi)\dagger \qquad (n = 0, 1, \cdots).$$

Hence, the denominators of the even convergents of (6) *coincide with the denominators of the convergents of* (5), which, therefore, can be obtained from the former by "contraction."

We wish now to find a simple expression for $B_{2n+1}(x)$ *in terms of the orthogonal and normal polynomials* (1). (8), with $\epsilon = 1$, is equivalent to

$$(12)\qquad \int_a^b B_{2n+1}(x)\,G_{n-1}(x)\,d\psi(x) = 0,$$

---

* O. Perron, ***Die Lehre von den Kettenbrüchen*** (1913), Ch. IX. This author considers $\int_a^b \frac{d\psi(y)}{x+y}$. It follows that the even $l-s$ in (6) differs in sign from the corresponding quantities used by Perron.

† Jacques Chokhate, "Sur le développement de l'intégrale $\int_a^b \frac{p(y)\,dy}{x-y}$ en fraction continue et sur les polynomes de Tchebycheff," *Rendiconti di Palermo*, Vol. 47 (1923), pp. 25-46.

which, combined with (7, 3, 4), yields:

$$(13)\quad B_{2n+1}(x) = l_1 l_2 \cdots l_{2n+1}\left[\Phi_{n+1}(x) - \Phi_n(x)\frac{\Phi_{n+1}(0)}{\Phi_n(0)}\right] \quad (n = 0, 1, \cdots),$$

i. e. $B_{2n+1}(x)$ is a linear combination of $\phi_n(x)$ and $\phi_{n+1}(x)$. This result is due to A. Markoff.*

It is our aim to derive from it a more interesting and important expression of $B_{2n+1}(x)$. For this we shall make use of Darboux's formulae, also of the polynomials $\phi_n(x; a, b; x d\psi)$.

2. Consider the relation †

$$(14)\quad \begin{aligned} G_n(x) &= \int_a^b K_n(x, t) G_n(t)\, d\psi(t) \\ K_n(x, t) &\equiv \sum_{i=0}^{n} \phi_i(x)\phi_i(t). \end{aligned}$$

Replacing here $G_n(t)$ by $(t-x)G_{n-1}(t)$, we get:

$$(15)\quad 0 = \int_a^b (t-x) K_n(x, t) G_{n-1}(t)\, d\psi(t),$$

which, combined with (2), gives:

$$(16)\quad (t-x)K_n(x, t) = A_{n+1}(x)\phi_{n+1}(t) + A_n(x)\phi_n(t).$$

Compare in (16) the coefficients of t^{n+1} and then interchange x and t:

$$A_{n+1}(x) = \frac{a_n}{a_{n+1}}\phi_n(x), \qquad A_n(x) = -\frac{a_n}{a_{n+1}}\phi_{n+1}(x).$$

Thus we get Darboux's formulae: ‡

$$(17)\quad K_n(x, t) = \frac{a_n}{a_{n+1}}\left[\frac{\phi_n(x)\phi_{n+1}(t) - \phi_n(t)\phi_{n+1}(x)}{t-x}\right].$$

$$(18)\quad \begin{aligned} K_n(x) \equiv K_n(x; d\psi) \equiv K_n(x, x) &\equiv \sum_{i=0}^{n} \phi_i^2(x) \\ &= \frac{a_n}{a_{n+1}}\left[\phi'_{n+1}(x)\phi_n(x) - \phi'_n(x)\phi_{n+1}(x)\right]. \end{aligned}$$

* A. Markoff, "Note sur les fractions continues," *Bulletin de l'Académie Impériale des Sciences de Russie*, (5), (1895), Vol. 2, pp. 9-16.

† J. Shohat, "On a General Formula in the Theory of Tchebycheff Polynomials and Its Applications," *Transactions of the American Mathematical Society*, Vol. 29 (1927), pp. 569-583.

‡ Darboux, "Mémoire sur l'approximation des fonctions de très grands nombres," *Journal des Mathématiques* (3), Vol. 4 (1878), pp. 5-56, 377-416. This process of deriving Darboux's formulae is due, in part, to M. J. Geronimus.

3. We assume now that

$$(19) \qquad x \leq a, \quad \text{or} \quad \geq b.$$

With such x we get readily, comparing (15) and (2):

$$K_n(x,t) = C_n(x)\phi_n(t;\ |t-x|\,d\psi),$$

and comparing the coefficients of t^n:

$$(20) \quad K_n(x,t) = [a_n(d\psi)/a_n(|t-x|\,d\psi)]\,\phi_n(x;\ d\psi)\cdot\phi_n(t;\ |t-x|\,d\psi) \qquad (x \geq a,\ \leq b)$$

$$(20\text{ bis}) \quad K_n(x) = [a_n(d\psi)/a_n(|t-x|\,d\psi)]\cdot\phi_n(x;\ d\psi)\cdot\phi_n(x;\ |t-x|\,d\psi)$$

$$(21) \quad K_n(0,t) = [a_n(d\psi)/a_n(td\psi)]\,\phi_n(0;\ d\psi)\phi_n(t;\ td\psi) \qquad (a \geq 0)$$

$$(21\text{ bis}) \quad K_n(0) = [a_n(d\psi)/a_n(td\psi)]\phi_n(0;\ d\psi)\phi_n(0;\ td\psi) = (b_{2n+2})^{1/2}\phi_n(0;\ d\psi)\phi_n(0;\ td\psi).$$

Formulae (20, 21) are of great importance in the theory of orthogonal Tchebycheff Polynomials.

4. *Hereafter* $a \geqq 0$. Making use of *

$$(22) \quad x\phi_n(x;\ xd\psi) = (b_{2n+3})^{1/2}\phi_{n+1}(x;\ d\psi) + (b_{2n+2})^{1/2}\phi_n(x;\ d\psi)$$

$$(23) \quad = -\phi_{n+1}(0;\ d\psi)/\phi_n(0;\ d\psi) = (b_{2n+2}/b_{2n+3})^{1/2}$$

$$(24) \quad l_{2n} = -1/b_{2n+1}l_{2n+1}; \quad \lambda_n = b_{2n-2}b_{2n-1} = a^2{}_{n-2}/a^2{}_{n-1}; \quad b_{2n+2} = a_n{}^2(d\psi)/a_n{}^2(xd\psi),$$

we get:

$$(25) \quad l_{2n} = \frac{1}{a^2{}_{n-1}\Phi_{n-1}(0)\Phi_n(0)} = \frac{1}{(\lambda_{n+1})^{1/2}\phi_{n-1}(0)\phi_n(0)}.$$

We put now $x = 0$ in (15) and compare with (12). Since $B_{2n+1}(0) = 0$, we have necessarily

$$(26) \quad \int_a^b [B_{2n+1}(t) - CtK_n(0,t)]\,G_{n-1}(t)\,d\psi(t) = 0,$$

where the constant C can be so chosen that $B_{2n+1}(t)/t - CK_n(0,t)$ becomes a polynomial *of degree* $n-1$. Identifying it with $G_{n-1}(t)$ in (26), we get, making use of (21, 24), *the fundamental result:*

$$B_{2n+1}(t)/t = CK_n(0,t) = D(b_{2n+2})^{1/2}\phi_n(0,\ d\psi)\phi_n(t;\ td\psi),$$

which gives, by virtue of (9, 23, 25), $D = 1$, and *we thus get the desired expression of* $B_{2n+1}(x)$ *in an unexpectedly simple form:*

* *Loc. cit.* (†).

$$(27)\qquad B_{2n+1}(x) \equiv (b_{2n+2})^{1/2} x\phi_n(0;\, d\psi)\,\phi_n(x;\, xd\psi) \equiv xK_n(0, x).$$

Hence, *the study of the denominators* $B_{2n+1}(x)$ *is reduced to that of* $\phi_n(x;\, xd\psi)$, which, for example, in the case of polynomials of Jacobi or Laguerre are orthogonal polynomials of the same class respectively.

5. We proceed to derive an application to mechanical quadratures. It is known * that if in the formula of mechanical quadratures

$$(28)\qquad \begin{gathered} \int_0^b f(x)\,d\psi(x) = \sum_{i=0}^{n} H_i f(x_i) + R_n(f) \\ 0 = x_0 < x_1 < \cdots < x_n < b; \qquad \phi(x) = \prod_{i=0}^{n} (x - x_i), \\ H_i = \int_0^b \frac{\phi(x)}{(x - x_i)\phi'(x_i)}\, d\psi(x) \end{gathered}$$

we wish to have $R_n(f) \equiv 0$ for $f(x) \equiv G_{2n}(x)$, we must choose $x_1, x_2, \cdots, x_n$ as zeros of $\phi_n(x;\, xd\psi)$. Formula (27) leads to the following

THEOREM. *The formula of mechanical quadratures* (28), *with* $n+1$ *ordinates, one of which is fixed at* $x_0 = 0$, *holds true* (i. e. $R_n \equiv 0$) *for an arbitrary polynomial of degree* $\leq 2n$, *if and only if the* $n+1$ *ordinates correspond to the zeros of the polynomial* $B_{2n+1}(x)$. Moreover, $H_0 = 1/K_n(0;\, d\psi)$.

The first part of our statement is proved by (27). In order to establish the second part, we write [see (21)]:

$$H_0 = \int_0^b \frac{x\phi_n(x;xd\psi)}{x(x\phi_n(x;xd\psi))'_{x=0}}\, d\psi(x) = \int_0^b \frac{\phi_n(x;xd\psi)}{\phi_n(0;xd\psi)}\, d\psi(x) =$$
$$= \int_0^b \frac{K_n(0;x)\,d\psi(x)}{K_n(0;d\psi)} = \frac{1}{K_n(0;d\psi)}.$$

This result is of importance in connection with the *practical application* of the formula (28).

We thus obtained in a general form a result previously established,† in an entirely different way, for a very special class of Jacobi polynomials.

As to the value of $K_n(0;\, d\psi)$, it can be found by means of (21 bis), or even still simpler in some special cases. As an illustration Laguerre polynomials may serve. Here

---

* J. Shohat, "On a Certain Formula of Mechanical Quadratures with Non-Equidistant Ordinates," *Transactions of the American Mathematical Society*, Vol. 31 (1930), pp. 448-463.

† *Loc. cit.* (*).

$$(a, b) = (0, \infty), \qquad d\psi(x) = e^{-x} dx,$$

$$K_n(0) = -e^{-x} K_n(x) \Big|_0^\infty = -\int_0^\infty [e^{-x} K_n(x)]' dx$$

$$= \int_0^\infty e^{-x} K_n(x)\, dx - \int_0^\infty e^{-x} K'_n(x)\, dx = n + 1.$$

[making use of (2)].

6. For the sake of simplicity we take $b = 1$. For n very large we ca derive, in a very simple manner, an upper limit for H_0, for the general cla of Tchebycheff polynomials corresponding to

(29) $$d\psi(x) = p(x)\, dx, \quad p(x) = x^\sigma q(x) \qquad (\sigma > -1,\ q(0) \neq 0),$$

$q(x)$ continuous at $x = 0$. We make use of the fact that *

(30) $$1/K_n(0) = \min. \int_0^1 G_n^2(x)\, d\psi(x), \quad \text{with } G_n(0) = 1.$$

Taking here $G_n(x) \equiv (1 - x)^n$, we get:

(31) $$H_0 = 1/K_n(0) < \int_0^1 (1-x)^{2n} p(x)\, dx = \int_0^1 x^{2n} p_1(x)\, dx$$

$$(p_1(x) \equiv p(1-x)).$$

On the other hand, (29) leads to †

(32) $$\int_0^1 x^{2n} p_1(x)\, dx = \frac{\Gamma(\sigma+1)\, q(0) + o(1)}{(2n)^{\sigma+1}} \qquad (n \to \infty).$$

Hence,

(33) $$H_0 = 1/K_n(0) = O(1/n^{\sigma+1}) \qquad (\text{under condition (29)}).$$

Remark. The above results hold, mutatis mutandis, for the upper end-point $(x = b)$ of the interval $(0, b)$ in question.

THE UNIVERSITY OF PENNSYLVANIA.

* *Loc. cit.* (†).

† J. Shohat, "On the Asymptotic Expressions of Certain Definite Integrals," *Annals of Mathematics*, Vol. 27 (1925), pp. 3-11.

ON A PROBLEM OF M. J. SHOHAT.

By J. GERONIMUS.

In his paper "On a General Formula in the Theory of Tchebycheff Polynomials and Its Applications," * M. J. Shohat considers the following problem: *Given an interval* (a, b) *finite or infinite, a function* $p(x)$ *non-negative on this interval, for which all integrals*

$$\int_a^b p(x)x^\nu dx, \qquad (\nu = 0, 1, 2, \cdots)$$

exist, find the minimum L *of the integral* $\int_a^b p(x)y^2(x)\,dx$, *where* $y(x)$ *is a polynomial of degree* $\leqq n$; $y(x) = \sum_{i=0}^{n} \sigma_i x^i$, *subject to the condition* $\omega(y) \equiv \sum_{i=0}^{n} \alpha_i \sigma_i = a$ $(a \neq 0)$ α_i *being given real numbers.*

In solving this problem M. Shohat considers certain polynomials which may be regarded as a generalization of orthogonal polynomials. It is the object of the present paper to generalize the problem of M. Shohat, assuming that the coefficients of our polynomial $y(x)$ are connected by $m+1$ linear relations. Some properties of these polynomials will be considered.

1. *Find the minimum value* L *of the integral*

$$(1) \qquad \int_a^b p(x)y^2(x)\,dx$$

under the following conditions:

$$(2) \qquad \omega_i(y) \equiv \sum_{s=0}^{n} \alpha_{is}[y^{(s)}(\eta)/s\,!] = d_i \qquad (i = 0, 1, 2, \cdots, m),$$

η, d_i *and* α_{is} $(i = 0, 1, 2, \cdots, m;\ s = 0, 1, 2, \cdots, n)$ *being given real numbers.*

Introduce $m+1$ polynomials $u_i(x)$ of the n-th degree possessing the property

$$(3) \qquad \int_a^b p(x)u_i(x)G_n(x)\,dx = \omega_i(G_n) \qquad (i = 0, 1, 2, \cdots, m),$$

* *Transactions of the American Mathematical Society*, Vol. 29 (1927), pp. 569-583.

where $G_n(x)$ is an *arbitrary* polynomial of degree $\leqq n$. For $\eta = 0$ we obtain the polynomials introduced by M. A. Angelesco and M. J. Shohat, which are a generalization of orthogonal polynomials.* Put

$$y(x) = \sum_{i=0}^{n} \gamma_i u_i(x) + z(x),$$

where γ_i are to be found from the conditions $\omega_k(z) = 0 \quad (k = 0, 1, 2, \cdots, m)$. It is always possible to determine the γ_i because the determinant of the above system of equations

$$a = \| a_{is} \| \qquad (i, s = 0, 1, 2, \cdots, m),$$

where

$$a_{is} = \int_a^b p(x) u_i(x) u_s(x)\, dx, \tag{4}$$

is positive, being the discriminant of the positive definite quadratic form

$$\sum_{i,s=0}^{m} a_{is}\gamma_i\gamma_s = \int_a^b p(x)\{\sum_{s=0}^{m} \gamma_s u_s(x)\}^2 dx.$$

We shall show that the condition of minimum of (1) *is:* $z(x) \equiv 0$. In fact, by virtue of the relations $\omega_i(z) = 0 \quad (i = 0, 1, \cdots, n)$,

$$\int_a^b p(x) y^2(x)\, dx = \int_a^b p(x) y_1^2(x)\, dx + \int_a^b p(x) z^2(x)\, dx \geqq \int_a^b p(x) y_1^2(x)\, dx$$

$$(y_1(x) = \sum_{s=0}^{m} \gamma_s u_s(x)),$$

which proves our statement.

Thus we see that the polynomial $y(x)$, for which the minimum of (1) is attained, may be found from the equation

$$\begin{vmatrix} y(x) & u_0(x) & u_1(x) & \cdots & u_m(x) \\ d_0 & a_{00} & a_{01} & \cdots & a_{0m} \\ d_1 & a_{10} & a_{11} & \cdots & a_{1m} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ d_m & a_{m0} & a_{m1} & \cdots & a_{mm} \end{vmatrix} = 0. \tag{5}$$

We find in a similar way that the minimal value L may be obtained from the equation:

* A. Angelesco, "Sur les polynomes orthogonaux en rapport avec d'autres polynomes," *Bulletin de la Société des Sciences de Cluj*, t. I (1921), pp. 44-59.

$$
\begin{vmatrix}
L & d_0 & d_1 & \cdots & d_m \\
d_0 & a_{00} & a_{01} & \cdots & a_{0m} \\
d_1 & a_{10} & a_{11} & \cdots & a_{1m} \\
\cdot & \cdot & \cdot & \cdots & \cdot \\
\cdot & \cdot & \cdot & \cdots & \cdot \\
d_m & a_{m0} & a_{m1} & \cdots & a_{mm}
\end{vmatrix} = 0. \tag{6}
$$

2. We can readily express $u_i(x)$ and a_{is} $(i, s = 0, 1, \cdots, m)$ in terms of orthogonal normal Tchebycheff polynomials $\phi_k(x)$ $(k = 0, 1, 2, \cdots)$, corresponding to the given characteristic function $p(x)$ and the interval (a, b). Using their fundamental property

$$
\int_a^b p(x)\phi_i(x)\phi_s(x)\,dx = 0\ (i \neq s),\ 1\ (i = s), \tag{7}
$$

we find easily from (3) and (4) that

$$
\begin{aligned}
u_i(x) &= \sum_{r=0}^{n} \omega_i(\phi_r)\phi_r(x), \\
a_{is} &= \sum_{r=0}^{n} \omega_i(\phi_r)\omega_s(\phi_r).
\end{aligned} \qquad (i, s = 0, 1, 2, \cdots, m). \tag{8}
$$

We can also give the solution of our problem in terms of the polynomials of Appell corresponding to the given characteristic function $p(x)$ and the interval (a, b).* Put

$$
c_k \equiv c_k(\eta) = \int_a^b p(x)(x - \eta)^k dx \qquad (k = 0, 1, 2, \cdots). \tag{9}
$$

Then our minimal problem is reduced to minimizing the quadratic form

$$
\sum_{i,s=0}^{n} a_i a_s c_{i+s}
$$

under the conditions

$$
\omega_k(y) = \sum_{s=0}^{n} \alpha_{ks} a_s = d_k, \qquad (k = 0, 1, 2, \cdots, m).
$$

Using the classical method we find readily:

---

* P. Appell, "Sur une classe de polynomes," *Annales de l'École Normale Supérieure*, t. 9 (1880), pp. 119-144.

$$(10)\quad \begin{vmatrix} y(x) & 1 & x-\eta & \cdots & (x-\eta)^n & 0 & 0 & \cdots & 0 \\ 0 & c_0 & c_1 & \cdots & c_n & \alpha_{00} & \alpha_{10} & \cdots & \alpha_{m0} \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & c_n & c_{n+1} & \cdots & c_{2n} & \alpha_{0n} & \alpha_{1n} & \cdots & \alpha_{mn} \\ d_0 & \alpha_{00} & \alpha_{01} & \cdots & \alpha_{0n} & 0 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdots & \cdot \\ d_m & \alpha_{m0} & \alpha_{m1} & \cdots & \alpha_{mn} & 0 & 0 & \cdots & 0 \end{vmatrix} = 0$$

$$(11)\quad \begin{vmatrix} L & 0 & 0 & \cdots & 0 & d_0 & d_1 & \cdots & d_m \\ 0 & c_0 & c_n & \cdots & c_n & \alpha_{00} & \alpha_{10} & \cdots & \alpha_{m0} \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & c_n & c_{n+1} & \cdots & c_{2n} & \alpha_{0n} & \alpha_{1n} & \cdots & \alpha_{mn} \\ d_0 & \alpha_{00} & \alpha_{01} & \cdots & \alpha_{0n} & 0 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdots & \cdot \\ d_m & \alpha_{m0} & \alpha_{m1} & \cdots & \alpha_{mn} & 0 & 0 & \cdots & 0 \end{vmatrix} = 0$$

To find an expression for the $u_i(x)$ and a_{is}, we put in (3)

$$G_n(x) = (x-\eta)^k \qquad (k = 0, 1, 2, \cdots, n),$$

which yields:

$$(12)\quad \begin{vmatrix} u_i(x) & 1 & x-\eta & \cdots & (x-\eta)^n \\ \alpha_{i0} & c_0 & c_1 & \cdots & c_n \\ \alpha_{i1} & c_1 & c_2 & \cdots & c_{n+1} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \alpha_{in} & c_n & c_{n+1} & \cdots & c_{2n} \end{vmatrix} = 0 \qquad (i = 0, 1, 2, \cdots, m),$$

$$(13)\quad \begin{vmatrix} a_{ik} & \alpha_{k0} & \alpha_{k1} & \cdots & \alpha_{kn} \\ \alpha_{i0} & c_0 & c_1 & \cdots & c_n \\ \alpha_{i1} & c_1 & c_2 & \cdots & c_{n+1} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \alpha_{in} & c_n & c_{n+1} & \cdots & c_{2n} \end{vmatrix} = 0 \qquad (i, k = 0, 1, \cdots, m).$$

3. Hereafter,

$$(14)\quad \omega_i(G_n) \equiv G_n^{(i)}(\eta) \qquad (i = 0, 1, 2, \cdots, m).$$

We see from (4) that

$$(15)\quad a_{ik} = \sum_{i=0}^{n} \phi_r^{(i)}(\eta)\,\phi_r^{(k)}(\eta) = (-1)^{i+k} i\,!\,k\,!\,C_n^{ik}/C_n \qquad (i, k = 0, 1, \cdots, m),$$

where $C_n = \| c_{ik} \|$ $(i, k = 0, 1, \cdots, n)$, $c_{ik} = c_{i+k}$ and C_n^{ik} is the minor of this determinant corresponding to the element c_{ik}. This formula leads to some interesting results involving persymmetric determinants formed by the polynomials of Appell.

$$(16)\quad \begin{vmatrix} H_1(z) & H_2(z) & \cdots & H_n(z) \\ H_2(z) & H_3(z) & \cdots & H_{n+1}(z) \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ H_n(z) & H_{n+1}(z) & \cdots & H_{2n-1}(z) \end{vmatrix} = \frac{2^{n(n-1)/2}1!\,2!\cdots(n-1)!}{i^{n^2}} \cdot H_n(iz),$$

where

$$(17)\qquad H_k(x) = (-1)^k e^{x^2}(d^k/dx^k)e^{-x^2} \qquad (k = 0, 1, 2, \cdots)$$

is the Hermite polynomial of degree k. Similarly,

$$(18)\quad \begin{vmatrix} P_1(z) & P_2(z) & \cdots & P_n(z) \\ P_2(z) & P_3(z) & \cdots & P_{n+1}(z) \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ P_n(z) & P_{n+1}(z) & \cdots & P_{2n-1}(z) \end{vmatrix} = \left(\frac{(z^2-1)^{n^2/2}}{2^{n^2-n}}\right) \cdot T_n\left(\frac{z}{(z^2-1)^{1/2}}\right),$$

where $P_k(z)$ is the Legendre polynomial and

$$(19)\qquad T_k(y) = \frac{[y + (y^2-1)^{1/2}]^k + [y - (y^2-1)^{1/2}]^k}{2}$$

is the trigonometric polynomial.*

Suppose further that $m = 0$ *and*

$$(20)\qquad \omega_0(G_n) \equiv \omega(G_n) \equiv G_n(c).$$

Then, putting

$$\phi_k(x) = \sum_{s=0}^{k} d_s^{(k)} x^{k-s} \qquad (k = 0, 1, 2, \cdots),$$

we see that the polynomial

$$(21)\quad u(x) \equiv \sum_{k=0}^{n} \phi_k(x)\phi_k(c) - \frac{d_0^{(n)}}{d_0^{(n+1)}} \cdot \frac{\phi_{n+1}(c)\phi_n(x) - \phi_n(c)\phi_{n+1}(x)}{c - x}$$

possesses the property

$$(22)\qquad \int_a^b p(x)u(x)G_n(x)\,dx = G_n(c).\dagger$$

* J. Geronimus, *On Some Persymmetric Determinants Formed by Polynomials of M. Appell.* (In press.)

† *Loc. cit.* (*).

Taking, for instance, $c=1$, $-a=b=1$, $G_n(x)=(x-z)^n$, we see that

$$(23)\qquad (1-z)^n=\int_{-1}^{1} p(x)(x-z)^n u(x)\,dx.^*$$

Returning to an arbitrary real value of c not belonging to the interval $a<x<b$, we see from (21) that $u(x)$ is identical with the orthogonal polynomial corresponding to the interval (a,b) and the characteristic function $p(x)(c-x)$. Hence, the following identity:

$$(24)\qquad \begin{vmatrix} c_0 & c_1 & \cdots & c_{n+1} \\ c_1 & c_2 & \cdots & c_{n+2} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ c_{n-1} & c_n & \cdots & c_{2n} \\ 1 & c & \cdots & c^{n+1} \\ 1 & x & \cdots & x^{n+1} \end{vmatrix} = \begin{vmatrix} c_0 & c_1 & \cdots & c_n & 1 \\ c_1 & c_2 & \cdots & c_{n+1} & c \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ c_n & c_{n+1} & \cdots & c_{2n} & c^n \\ 1 & x & \cdots & x^n & 0 \end{vmatrix} \cdot (c-x),$$

where the c's are arbitrary numbers.

4. Suppose that

$$a_{ik}=\sum_{s=0}^{n}\phi_s^{(i)}(\eta)\phi_s^{(k)}(\eta), \qquad (i,k=0,1,\cdots,m)$$

is expressible as

$$(25)\qquad a_{ik}=\alpha_i^{(n)}\alpha_k^{(n)}h_{ik},$$

where $\alpha_i^{(n)}$ depends on n and i and h_{ik} depends on i and k only. Such is the case, for instance, for Legendre polynomials,† where

$$(26)\quad a_{ik} = \frac{(n+i+1)(n+k+1)P_n^{(i)}(\eta)P_n^{(k)}(\eta)+(1-\eta^2)P_n^{(i+1)}(\eta)P_n^{(k+1)}(\eta)}{2(i+k+1)},$$

if $\eta=+1,-1,0$. Under the condition (25), we have

$$(27)\qquad \frac{A_m^{ik}}{A_m}=\frac{1}{\alpha_i^{(n)}\alpha_k^{(n)}}\cdot\frac{H_m^{ik}}{H_m},$$

where

$$A_m=\|a_{ik}\|,\qquad H_m=\|h_{ik}\| \qquad (i,k=0,1,2,\cdots,m),$$

* *Loc. cit.* (†).

† J. Geronimus, "Sur l'écart moyen quadratique minimal de zéro d'un polynome dans l'intervalle donné" (in Russian), *Transactions of the Kharkow Mathematical Society*, Série 4, t. 2 (1928), p. 17.

and A_m^{ik}, H_m^{ik} resp. are the corresponding minors of these determinants. In a similar manner we may write

$$\frac{\bar{A}_m^{ik}}{\bar{A}_m} = \frac{1}{\alpha_i^{(m)}\alpha_k^{(m)}} \cdot \frac{H_m^{ik}}{H_m}, \tag{28}$$

where $\bar{A}_m = \| \bar{a}_{ik} \|$ $(i, k = 0, 1, \cdots, m)$ and

$$\bar{a}_{ik} = \sum_{r=0}^{m} \phi_r^{(i)}(\eta)\phi_r^{(k)}(\eta). \tag{29}$$

Thus we obtain:

$$\frac{A_m^{ik}}{A_m} = \frac{\alpha_i^{(m)}\alpha_k^{(m)}}{\alpha_i^{(n)}\alpha_k^{(n)}} \cdot \frac{\bar{A}_m^{ik}}{\bar{A}_m}.$$

Using (15) and the properties of minors of reciprocal determinants,* we obtain the following formula:

$$\frac{\bar{A}_m^{ik}}{\bar{A}_m} = (-1)^{i+k} \cdot \frac{c_{ik}}{i!\,k!},$$

and finally,

$$\frac{A_m^{ik}}{A_m} = \frac{\alpha_i^{(m)}\alpha_k^{(m)}}{\alpha_i^{(n)}\alpha_k^{(n)}} \cdot \frac{(-1)^{i+k}}{i!\,k!}\, c_{ik} \qquad (i, k = 0, 1, \cdots, m). \tag{30}$$

Returning to our minimal problem and using (6), we obtain the following result.

The minimal value L of the integral $\int_a^b p(x)y^2(x)\,dx$, being given $y^{(s)}(\eta) = d_s$ $(s = 0, 1, 2, \cdots, m)$, may be written thus [*when* (25) *holds true*]: $L = \int_a^b p(x)y_1^2(x)\,dx$, $y_1(x) = \sum_{i=0}^{m} \frac{\alpha_i^{(m)}}{\alpha_i^{(n)}} \cdot \frac{(x-\eta)^i}{i!}\, d_i$.

Example. $p(x) = 1$, $-a = b = 1$, $\eta = \pm 1$. Here

$$y_1(x) = \sum_{i=0}^{m} \frac{(m+i+1)P_m^{(i)}(\eta)}{(n+i+1)P_n^{(i)}(\eta)} \cdot \frac{(x-\eta)^i}{i!}\, d_i.$$

The same formula is valid for $\eta = 0$, but in this case we must take $m - 1$ (resp. $n - 1$) instead of m (resp. n), when $m - i$ (resp. $n - i$) is odd.

---

* E. Pascal, *I determinanti* (1923), p. 52.

ON THE CRITICAL POINTS OF NON-DEGENERATE NEWTONIAN POTENTIALS.

By Tsai-Han Kiang.

I. Introduction.

1. *Definitions.** In this paper an application will be made of a theorem of Morse on the relations between the critical points of a real function of n independent variables (Theorem 5, Morse I, pp. 387-388) to the proof of the existence of critical points, or equilibrium points as better known in physics, of a class of Newtonian potentials.†

We shall use the terminology of analysis situs as defined by Alexander,‡ substituting the terms *cycle* and *Betti number* for *closed chain* and *connectivity number* respectively. Terms in analysis situs not found in Alexander's paper will be as defined in Lefschetz's *Colloquium Lectures on Topology.* Throughout the present paper the terms *Betti number* and *manifold* will be used as abbreviations of *Betti number, mod* 2 and *orientable manifold* respectively.

Let $f(x)$ be a real single-valued function of three variables $(x) = (x_1,$

* The following works will be cited repeatedly in this paper. Hereafter each of these works will be referred to by the name of its author followed by a Roman numeral when necessary.

Kellogg: O. D. Kellogg, *Foundations of Potential Theory*, Berlin (1929); Morse I: M. Morse, "Relations Between the Critical Points of a Real Function of n Independent Variables," *Transactions of American Mathematical Society*, Vol. 30 (1925), pp. 345-396; Morse II: M. Morse, "The Analysis and Analysis Situs of Regular n-Spreads in $(n+r)$-space," *Proceedings of the National Academy of Sciences*, Vol. 12 (1927), pp. 813-817.

† The method of the present paper has been extended by the author in another paper, "On the Existence of Critical Points of Green's Functions for Three-Dimensional Regions," which will be published shortly elsewhere. In the latter the regions considered are of very general types from the point of view of analysis situs. Both of these papers are parts of the author's thesis at Harvard University, accepted in June, 1930. Since they were practically completed, there has appeared a paper by J. J. Gergen, "Mapping of a General Type of Three-Dimensional Region on a Sphere," *American Journal of Mathematics*, Vol. 52 (1930), pp. 197-224. In his paper Gergen showed that any Green's function for a three-dimensional region bounded by a torus has a critical point.

‡ J. W. Alexander, "Combinatorial Analysis Situs," *Transactions of the American Mathematical Society*, Vol. 28 (1926), pp. 301-329.

x_2, x_3). A point will be called a *critical point* of the function f if it is a zero of all the three partial derivatives of the first order of f. The value the function f assumes at a critical point will be called a *critical value* of f. A critical point will be called *degenerate* or *non-degenerate* according as it is or is not a zero of the hessian of the function. A function without degenerate critical points will be called a *non-degenerate* function.

Suppose f is of class C^2 * in a neighborhood of a non-degenerate critical point (x^0). For simplicity, we suppose $(x^0) = (0)$. Let f^0_{ij} be the partial derivative of f of second order with respect to x_i and x_j evaluated at (0). By a real non-singular linear transformation the non-singular quadratic form $f^0_{ij}x_ix_j$ † can be reduced to one of the forms:

$$\pm x_1^2 \pm x_2^2 \pm x_3^2.$$

The number of negative signs is called the *type number* of the critical point (0) of f by Morse in Morse I. There are four types of critical points.

2. *Morse's theorem.* Suppose a real, single-valued, non-degenerate function $f(x)$ is defined in a closed bounded region R of the real 3-space of three real variables $(x) = (x_1, x_2, x_3)$, and the function $f(x)$ is of class C^2 in the interior of R. We shall assume that f in R fulfills the following boundary conditions.

(1) The partial derivatives of f of the first and second orders take on continuous boundary values on the boundary B of R.

(2) The boundary B of R consists of two sets, B' and B'', of closed regular surfaces of class C^2.‡ The value of f on all the surfaces of B' is equal to a constant c' and the value of f on all the surfaces of B'' is equal to a constant c'', where c' is greater but c'' is less than the value of f at any interior point of R.

---

* A function is said to be of class C^n, if it is continuous together with all its partial derivatives of orders $1, 2, \cdots, n$. The hypothesis of C^3 of f in a neighborhood of the critical point in Morse I has been shown since by Morse to be unnecessary.

† We adopt the convention in tensor analysis under which a repetition of a subscript in a product means that the product is to be summed for admissible values of the subscript.

‡ A regular surface is a closed, bounded, and connected set of points the coördinates of points neighboring any given point of which can be represented by three functions of class C^1 of two parameters, in such a way that the jacobians of the functions with respect to the two parameters do not vanish simultaneously. If, moreover, the functions are of class C^n, $n > 1$, the surface is said to be of class C^n.

(3) At any point of B the value of the outer normal derivative f_n of f with respect to R * does not vanish.

Under all the above conditions the function f will be said to be *an admissible function in* R.

The region R can be shown to be a 3-complex.† From (2) and (3), at any point of B' the value of the outer normal derivative f_n of f is positive and at any point of B'' the value of f_n is negative. A surface of B' or of B'', as characterized by this statement, will be called *a surface of positive or negative type with respect to* f *and* R.

From a discussion of the trajectories orthogonal to the manifolds $f =$ constant one can show that the region R' of points in R satisfying

$$f \leqq c'' + e^2,$$

e being a sufficiently small positive number, has the same Betti numbers as the manifold $f = c''$. See Morse I, pp. 355-360. From Theorem 5, Morse I, pp. 387-388, as applied to the present regions R' and R, we have the following theorem.

THEOREM A. *Let* $f(x)$ *be an admissible function in a region* R. *Let* M_k $(k = 0, 1, 2, 3)$ *be the number of critical points of the kth type in the region. Let* R_k *and* R_k' *be the kth Betti numbers of the complexes* R *and* B'' *respectively. Then there exist integers* M_k^+ *and* M_k^- *of which* M_3^+ *and* M_0^- *are zero and the rest are positive or zero such that*

$$M_k = M_k^+ + M_k^-, \quad (k = 0, 1, 2, 3), \tag{2.1}$$

and such that

$$R_k - R_k' = M_k^+ - M_{k+1}^-, \quad (k = 0, 1, 2). \tag{2.2}$$

3. *New boundary conditions.* In our investigation we do not meet with exactly the boundary conditions of § 2, but the following ones.

(1a) The boundary B of R consists of two sets, B_1 and B_2, of closed regular surfaces. Each surface of the set B_1 is of class C^2. On such a surface the value of f is equal to a constant and the partial derivatives of f of the first and second orders take on continuous boundary values.

* By the outer normal derivative f_n at a point of B with respect to R we mean the unilateral directional derivative of f at that point along the normal to B in the sense that leads from interior points of R to the boundary.

† S. S. Cairns, "The Cellular Structure and Approximation of Regular Spreads," *Proceedings of the National Acaedmy of Sciences*, Vol. 16 (1930), pp. 488-491.

(2a) Each surface of the set B_2 is of class C^3. On such a surface the partial derivatives of f of the first order take on continuous boundary values.

(3a) The boundary surfaces may also be groupped into two sets, B' and B'', such that every surface of B' is of positive type with respect to f and R and every surface of B'' is of negative type (§ 2).

Under this set of new boundary conditions instead of that of § 2, the function f will be called *an admissible function in R under the new boundary conditions.*

These new boundary conditions also imply that the region R is a 3-complex and that f has no critical point on B.

On comparing this set of new boundary conditions with that set of § 2, one finds that an admissible function in R under the new boundary conditions with B_2 as an empty set of points differs from an admissible function in R of § 2 in that the former does not necessarily take on the absolute maximum on and only on all the boundary surfaces of positive type and the absolute minimum on and only on all the boundary surfaces of negative type. However we shall prove the following theorem.

THEOREM B. *Theorem A holds for an admissible function f in R under the new boundary conditions.*

The proof of this theorem consists in the reduction of the new boundary conditions to those of § 2. The reduction will be carried out in the following two steps in accordance with the remark on the difference of the two sets of boundary conditions made above.

First step. It can be shown that, by a redefinition of f in R in a neighborhood of each surface of B_2 (cf. Morse I, pp. 394-396), we shall obtain a new function $\bar{f}$ in a new region $\bar{R}$ which fulfills the new boundary conditions with $\bar{B}_2$ * as an empty set of points and which has the following properties: (a) The surfaces B_1 are still among the boundary surfaces of $\bar{R}$, and $\bar{R}$ is a closed subregion of R, homeomorphic to R. (b) The two sets, $\bar{B}'$ and $\bar{B}''$, of boundary surfaces of positive and negative types with respect to $\bar{f}$ and $\bar{R}$ are respectively homeomorphic to the two sets, B' and B'', of boundary surfaces of positive and negative types with respect to f and R. (c) The critical points of $\bar{f}$ are identical in position and type with those of f.

Second step. Now, suppose a function f in R fulfills the new boundary

* The set $\bar{B}_2$ for $\bar{f}$ and $\bar{R}$ is taken to have the same meaning as the set B_2 for f and R. This convention will be followed in this section.

conditions with B_2 as an empty set of points. We shall show that by a redefinition of f in R in a neighborhood of B in a similar but simpler way, we can obtain an admissible function $\bar{f}$ in the same region R in the sense of § 2, with the following properties: (a_1) A boundary surface of B of positive or negative type with respect to f and R is of the same type with respect to $\bar{f}$ and R. (b_1) The critical points of $\bar{f}$ are identical in position and type with those of f.

Let us fix our attention on one surface S of B', on which $f = d$, say. Since no critical point of f is on S, through each point of S there is one and only one orthogonal trajectory. We can represent each orthogonal trajectory in such a way that the parameter t at any point equals the value of f at that point (Morse I, p. 358). The parameter t for each trajectory is equal to d on S. It is possible to choose a constant $t_1 < d$ so that the points on all the trajectories for which $t_1 \leqq t \leqq d$ constitute a neighborhood N of S, in which there are no critical points of f and hence through any point of N there is one and only one such trajectory. Let t_2 be another constant between t_1 and d. For those points in N for which $t_1 \leqq t \leqq t_2$, set

$$\bar{f}(x) = f(x),$$

and for those points in N for which $t_2 \leqq t \leqq d$, set

$$\bar{f}(x) = f(x) + K(f(x) - t_2)^4,$$

where K is a sufficiently large positive constant. Then we find at points in N for which $t_2 \leqq t \leqq d$,

$$\bar{f}_i = f_i(1 + 4K(f - t_2)^3),$$

$$\bar{f}_{ij} = f_{ij}(1 + 4K(f - t_2)^3) + f_i f_j \cdot 12K(f - t_2)^2.$$

Hence, in the interior of R the function $\bar{f}$ is still of class C^2, and on S its partial derivatives of the first and second orders take on continuous boundary values. From the relations between $\bar{f}_i$ and f_i and the fact that $4K(f - t_2)^3 \geqq 0$ at points in N for which $t_2 \leqq t \leqq d$, $\bar{f}$ has no critical point in N. On S,

$$\bar{f} = d + K(d - t_2)^4,$$

an arbitrarily large constant. Then S is a surface of positive type with respect to $\bar{f}$ and R.

Similar argument can be applied to any boundary surface T of the set B''. In this case we shall obtain a new function with the same properties

except that this new function equals an arbitrarily small constant on T and that T is a surface of negative type with respect to this new function and R.

If we redefine f in neighborhoods of all the boundary surfaces simultaneously, and choose a suitable constant K for each surface in the redefinition, the new function obtained is admissible in R in the sense of § 2, has the property (a_1), and has its critical points identical in position with those of f. Since the type of a critical point is defined with reference to a small neighborhood of the critical point, and since the new function and the original function f are identical in a small neighborhood of each critical point, the new function has the property (b_1) also.

This completes the reduction of the new boundary conditions to those of § 2. Hence our theorem is proved.

II. The Newtonian Potential Due to a Non-Degenerate Admissible Volume, Simple, or Double Distribution.

4. *The potential.* Suppose in the 3-space there are s distinct closed regular surfaces $E_k (k = 1, 2, \cdots, s)$ of class C^3, no two of which enclose a common point. Let D designate the closed infinite region bounded by these surfaces and let E designate the sum of the surfaces E_k.

We shall consider a point at infinity added to the 3-space, so that the extended space becomes homeomorphic to the 3-sphere.

Suppose there is a volume distribution of matter of piecewise continuous density in the closed finite regions bounded by E, or a simple distribution of matter of density satisfying uniformly a Hölder's condition (Kellogg, p. 152) over E_k, or a double distribution of moment of class C^3 over E_k. In the case of double distribution, it should be understood that the potential and its partial derivatives of the first order at any point P of E are defined by their limiting values at P as P is approached from the interior points of D. The potential due to the volume, simple, or double distribution is therefore analytic in the finite part of $D - E$, and of class C^1 in the finite part of D (Kellogg, pp. 150-172).

Definition 1. *The distribution will be termed admissible, if on E the outer normal derivative f_n of the potential f with respect to D never vanishes. If moreover the potential f has no degenerate critical point in the finite part of D, the distribution will be termed non-degenerate.**

* The restriction that f be non-degenerate and admissible could be removed by the aid of the more general theorems on critical points in Morse II and in a recent paper by Morse, "The Critical Points of a Function of n Variables," *Transactions of*

At great distances, we have the following developments in spherical harmonics for f and its three partial derivatives of the first order $f_i (i = 1, 2, 3)$:

$$
(4.1) \qquad \begin{aligned} f(x) &= m/r + h_1(x)/r^3 + \cdots, \\ f_i(x) &= -mx_i/r^3 + h_2{}^i(x)/r^5 + \cdots, \end{aligned}
$$

where m is the total mass of the distribution ($m = 0$ in the case of a double distribution), r the distance from the origin to the variable point (x), and h's are the spherical harmonics of the orders indicated by their subscripts (Kellogg, pp. 143-145).

At the point at infinity, the potential f and its partial derivatives will be assigned the value zero, their limiting value. They are regular at the point at infinity (Kellogg, p. 217), and harmonic in D (Kellogg, p. 211). The values of f_n on each boundary surface E_k are of the same sign. We suppose the surfaces E_i $(i = 1, 2, \cdots, p)$ are of positive type with respect to f and D (§ 2), and the surfaces E_j $(j = p+1, p+2, \cdots, p+n)$ are of negative type, where $p + n = s$.

5. *Remarks on the critical points of f.* It is a well-known property of harmonic functions that the function f has neither maximum nor minimum in $D - E$; in other words, f has neither critical point of type 0 nor critical point of type 3 in $D - E$.

Suppose $h(x)$ is a harmonic function, but not a constant, in an open region H, finite or infinite. Let H' be a closed regular subregion * of H. If H' is finite, the critical points of h in H' constitute at most a finite number of isolated points, analytic curves, and analytic surfaces, † but can not fill any regular surface element, and the critical values of $h(x)$ in H' are finite in num-

---

the American Mathematical Society, Vol. 37 (1931), pp. 72-91. This would introduce a complexity which it seems desirable to avoid in a first treatment.

* A region is said to be regular, if it is a set of points in the extended 3-space and if its complete boundary consists of a finite number of closed regular surfaces.

† *A real analytic curve* is a closed, bounded, and connected set of points the coördinates (x_1, x_2, x_3) of points neighboring a given point of which are representable as real analytic functions of a suitable real parameter t, will not all three functions identically constant. *A real analytic surface* is a closed, bounded, and connected set of points defined by setting one or more real analytic functions of (x_1, x_2, x_3) equal to zero, such that the equations are not satisfied identically in the neighborhood of any point of the set and such that the points neighboring any point P of the set are neither included on an analytic curve nor reduce to the point P.

ber (Kellogg, p. 262, pp. 276-277). If H' is infinite, these statements can be proved as follows.*

In a spherical neighborhood N of the point at infinity the function $h(x)$ and its partial derivatives of the first order $h_i(x)$, $(i = 1, 2, 3)$, are harmonic. Let us subject these four functions to an inversion in the unit sphere with center at the origin taken at a point not in H':

$$x_i = x_i'/r'^2. \tag{5.1}$$

Let us designate the four transformed functions divided by r' by $F(x')$ and $F_i(x')$ respectively. After the singularity at the origin is removed, each of these functions is harmonic in the transformed neighborhood N' of the origin (Kellogg, pp. 232-233). These functions are analytic in N', and vanish at the transformed origin. Hence, in N', the locus of their common zeros consists of at most a finite number of isolated points, analytic curves, and analytic surfaces. If we note that the critical points of h in N are the common zeros of the four functions $F(x')$ and $F_i(x')$ in N' and conversely, what is to be proved follows immediately.

The author has a proof for the fact that, of a set of points constituting a real analytic surface, there is always a subset of points constituting a regular surface element.† But the nonexistence of a regular surface element of critical points has been proved by Kellogg (Kellogg, p. 262). Hence we have the following theorem.

THEOREM 1. *If $h(x)$ is a function harmonic in an infinite open region of the extended space but not a constant, then in any closed regular subregion, the locus of the critical points of h consists of at most a finite number of isolated points and analytic curves, and the number of critical values of h is finite.*

COROLLARY. *In the region D of § 4 the locus of the critical points of the non-degenerate admissible potential f consists of at most a finite number of isolated points, and the number of critical values of f is finite.*

6. *Nature of the locus at the point at infinity.* To study the locus $f(x) = 0$ in the neighborhood of the point at infinity, we invert it into a bounded locus. Suppose the origin is taken at a point at which there is no

---

* If a set of points in a regular infinite region is transformed by an inversion into a real analytic curve or surface, the set will be called a real analytic curve or surface.

† This will be published in a note elsewhere.

distribution and which is not on the locus $f = 0$. Then by the inversion (5.1) the potential f is transformed into the function

$$g(x') \equiv r' \cdot F(x'), \tag{6.1}$$

where F, upon removing a singularity at the origin, becomes harmonic in a neighborhood of the origin (Kellogg, p. 232). From the expression for f in (4.1), it is obvious that $F(0) = m$, where m is the total mass of the distribution.

If m is not zero, in a sufficiently small neighborhood of the origin F will not vanish, and will have the same sign as m, and so will g.

If m is zero, F being harmonic, will be positive at some points and negative at some points in any neighborhood of the origin, and so will g. Thus we have the following results.

LEMMA 1. *Suppose f is the potential due to an admissible distribution, the total mass of which is not zero. Then the point at infinity will be a point of proper minimum or maximum of f according as the total mass is positive or negative.*

If the total mass is zero, the point at infinity is not an isolated point on $f = 0$, and not a point of maximum or minimum of f.

7. *Equipotential surfaces in the neighborhood of the point at infinity.* $m \neq 0$. We shall call a regular surface homeomorphic to an ordinary sphere a *sphere-like surface.*

We shall prove the following lemma.

LEMMA 2. *For a sufficiently small positive constant e we have the following:*

(1) *If the total mass of the distribution is positive [negative], the equipotential locus $f = e^2$ $[f = -e^2]$ consists in part at least of an analytic sphere-like surface $G[H]$ enclosing all the boundary surfaces $E_k (k = 1, 2, \cdots, s)$.*

(2) *The infinite closed region bounded by $G[H]$ is homeomorphic to a 3-cell.*

(3) *The function f has no critical point in the finite part of the infinite closed region bounded by $G[H]$.*

(4) *At any point of $G[H]$ the outer normal derivative f_n of f with respect to the finite region bounded by $G[H]$ is negative [positive].*

Proof. In a neighborhood of the point at infinity the equipotential surfaces of the potential f due to any admissible distribution are transformed by

the inversion (5.1) into the equipotential surfaces of $g(x') \equiv r' \cdot F(x')$ in the corresponding neighborhood N of the origin. Subject $g(x')$ to the transformation

$$\bar{x} = F(x') \cdot x_i', \quad (i = 1, 2, 3). \tag{7.1}$$

The jacobian of this transformation at the origin is equal to m^3, which is not zero. The transformation is one-to-one and continuous in the neighborhood N, provided N is sufficiently small.

Suppose m is positive. Let e be a sufficiently small positive constant and $\bar{r}$ the distance from the origin to $(\bar{x}_1, \bar{x}_2, \bar{x}_3)$. The locus of points satisfying $g(x') \leqq e^2$ in N will be carried by (7.1) into the locus of points satisfying $\bar{r} \leqq e^2$, a spherical region. The function $\bar{r}(\bar{x}_1, \bar{x}_2, \bar{x}_3)$ has no critical point for which $0 < \bar{r} \leqq e^2$. Thus (1), (2), and (3) of the lemma are proved for m positive. The proof is similar for m negative.

Since f has no critical point on $G[H]$ and since f_n is normal to $G[H]$, f_n can never vanish on $G[H]$. This proves (4) of the lemma.

The surface $G[H]$ of Lemma 2 will be called *the outer branch of* $f = e^2 [f = -e^2]$.

8. *Critical points of* f. $m \neq 0$. Suppose f is the potential due to a non-degenerate admissible distribution. There are two cases.

Case 1. $m < 0$. The outer branch H of $f = -e^2$ of Lemma 2 together with boundary surfaces $E_1, E_2, \cdots, E_{p+n}$ of D (§ 4) bounds a closed bounded region which we take as the region R of § 3. The boundary of R we have designated by B. The set B' of boundary surfaces of positive type with respect to f and R (§ 3) consists of $p + 1$ surfaces, $H, E_1, E_2, \cdots, E_p$, and the set B'' of boundary surfaces of negative type with respect to f and R consists of the remaining n surfaces, $E_{p+1}, E_{p+2}, \cdots, E_{p+n}$. In the region R, the function f is admissible under the new boundary conditions of § 3.

Let the first Betti number of E_k be $2g_k$. Then the first Betti number of either one of the two regions into which the extended 3-space (§ 4) is divided by E_k is g_k (Morse II, Corollary to Theorem 9 and the second part of Theorem 11, pp. 816-817). The Betti numbers of the set B'' of boundary surfaces are the following:

$$R'_0 = n, \; R'_1 = 2 \sum_{j=p+1}^{p+n} g_j, \; R'_2 = n, \; R'_3 = 0.$$

Since no two of the surfaces E_k enclose a common point, the Betti numbers of the region complimentary to R with respect to the extended 3-space are therefore

$$p+n+1,\ \sum_{k=1}^{p+n} g_k,\ 0,\ 0.$$

From another theorem in analysis situs (Morse II, Corollary to Theorem 11, p. 817), the Betti numbers of R are the following:

$$R_0=1,\ R_1=\sum_{k=1}^{p+n} g_k,\ R_2=p+n,\ R_3=0.$$

The function f is harmonic in R and therefore in the interior of R it has neither critical point of type 0 nor critical point of type 3, that is, it has M_0^+ and M_3^- both equal to zero. From Theorem B and (2.2), we have for f in R

$$M_1^-=n-1, \qquad M_1^+=\sum_{i=1}^{p} g_i+b,$$

$$M_2^-=\sum_{j=p+1}^{p+n} g_j+b,\ M_2^+=p,$$

where b is an integer, positive, negative, or zero, but not less than the larger one of the two numbers:

$$-\sum_{i=1}^{p} g_i,\quad -\sum_{j=p+1}^{p+n} g_j.$$

From Theorem B and (2.1) we have for f in R

$$M_1=n-1+\sum_{i=1}^{p} g_i+b,$$

(8.1)

$$M_2=p\qquad+\sum_{j=p+1}^{p+n} g_j+b.$$

Case 2. $m>0$. The outer branch G of $f=e^2$ of Lemma 2 together with the boundary surfaces $E_1, E_2,\cdots,E_{p+n}$ bounds a closed bounded region which we take as the region R. The boundary of R we have designated by B. The set B' of boundary surfaces of positive type with respect to f and R consists of p surfaces, $E_1, E_2,\cdots,E_p$, and the set B'' of boundary surfaces of negative type with respect to f and R consists of the remaining $n+1$ surfaces, $G, E_{p+1}, E_{p+2},\cdots,E_{p+n}$. In the region R the function f is admissible under the new boundary conditions of § 3. As in the preceding case, we find for f in R

$$M_1^-=n,\quad M_1^+=\sum_{i=1}^{p} g_i+b,$$

$$M_2^-=\sum_{j=p+1}^{p+n} g_j+b,\quad M_2^+=p-1;$$

and

$$M_1=n+\sum_{i=1}^{p} g_i+b,$$

(8.2)

$$M_2=p-1+\sum_{j=p+1}^{p+n} g_j+b;$$

where b has the same meaning as in the preceding case. Hence we have the following theorem.

THEOREM 2. *Suppose there is a non-degenerate admissible distribution over, or in the closed finite regions bounded by, $p+n$ distinct closed regular surfaces, $E_1, E_2, \cdots, E_{p+n}$, of class C^3, no two of these surfaces enclosing a common point. Suppose the first Betti number of $E_k (k=1,2,\cdots,p+n)$ is $2g_k$. Suppose $E_i (i=1,2,\cdots,p)$ is of positive type with respect to f and the space exterior to it, and $E_j (j=p+1, p+2, \cdots, p+n)$ is of negative type with respect to f and the space exterior to it* (§ 2).

Then, if the total mass of the distribution is not equal to zero, in the space exterior to all these $p+n$ surfaces the potential f due to this distribution has at least

$$p+n-1+\left|\sum_{i=1}^{p} g_i - \sum_{j=p+1}^{p+n} g_j\right|$$

critical points. If f has more than this number of critical points, the excess is even.

According as the total mass is negative or positive, the respective numbers M_1 and M_2 of critical points of type 1 *and type* 2 *are given by* (8.1) *or* (8.2).

9. *Equipotential surfaces of a redefined function g in the neighborhood of the infinite branch of $f=0$. $m=0$.* The function f for $m=0$ in the region D of § 4 fulfills the conditions for an admissible function in D under the new boundary conditions of § 3 except that D is not bounded. But the argument for redefinition of a function in R in a neighborhood of B'' as described in § 3 can still be applied to redefinition of f in D in a neighborhood of the boundary E of D. Let the new function obtained from such a redefinition of f in D in a neighborhood of E be g. We shall still designate the new region by D and its boundary by E. The function g in the new region D fulfills the conditions for an admissible function in D in the sense of § 2 except that D is not bounded.

By an inversion with the center of inversion not on the locus $g=0$, the locus $g=0$ will be transformed into a bounded locus through the origin. We shall say that a point of $g=0$ is *connected to the point at infinity* if it is transformed by the inversion into a point which can be connected to the origin on the transformed locus. We shall call the locus of points of $g=0$ which are connected to the point at infinity the *infinite branch* of $g=0$. Similarly the infinite branch of $f=0$ will be defined.

Let us suppose f is non-degenerate. From the Corollary to Theorem 1, in D the number of critical values of f, and therefore of g, is finite. In

redefinition we can, and shall, assume that the boundary values c' and c'' (§ 2) of g are both different from zero. It is then possible to take a positive constant e so small that no critical value of g lies between 0 and $+e^2$ or between $-e^2$ and 0, that $c'' < -e^2 < +e^2 < c'$, and that $\pm e^2$ are not critical values of g. The open set of points satisfying $-e^2 < g < e^2$ contains the point at infinity. Let us suppose the above center of inversion is not in this set. Then this set will be transformed into an open bounded set which consists of one or more continua. As a convention we shall say that the subset of points satisfying $-e^2 < g < e^2$, which is transformed into the continuum containing the origin, is a continuum, and we shall call it the *infinite continuum*. In the infinite continuum the infinite branch of $g = 0$ is contained but not the boundary E. The boundary of the infinite continuum consists of $u \geqq 1$ closed regular surfaces S_α $(\alpha = 1, 2, \cdots, u)$ of class C^2 satisfying $g = -e^2$, and $v \geqq 1$ closed regular surfaces S_β $(\beta = u+1, u+2, \cdots, u+v)$ of class C^2 satisfying $g = e^2$. S_α will be called the u outer branches of $g = -e^2$ and S_β the v outer branches of $g = e^2$. Each of these $u + v$ surfaces is a 2-manifold. No two of these surfaces enclose a common point. The sum of the finite closed regions bounded by these surfaces and the infinite continuum is the extended 3-space. It follows from a method of proof given by Morse (Morse I, pp. 356-360) that the Betti Numbers of each of these finite regions are independent of the constant e, for e sufficiently small and positive.

Since the value of g is $-e^2$ on S_α and e^2 on S_β, and since $\pm e^2$ are not critical values, on each of the surfaces S_α and S_β, the outer normal derivative of g with respect to the finite space it bounds never vanishes. It is then easy to see that at each point of S_α the outer normal derivative of g is positive, but at each point of S_β the outer normal derivative of g is negative. Moreover any critical point of g not on the infinite branch of $g = 0$ is enclosed by some one of the surfaces S_α and S_β.

These results will be summed up in the following lemma.

LEMMA 3. *Suppose f is the potential due to a non-degenerate admissible distribution, the total mass of which is zero. Suppose g is the function obtained by redefining f in a neighborhood of E (in the way of redefining f in a neighborhood of B'' as described in* § 3). *For a sufficiently small positive constant e, there are $u \geqq 1$ closed regular surfaces S_α $(\alpha = 1, 2, \cdots, u)$ of class C^2 satisfying $g = -e^2$, the u outer branches of $g = -e^2$, and $v \geqq 1$ closed regular surfaces S_β $(\beta = u+1, u+2, \cdots, u+v)$ of class C^2 satisfying $g = e^2$, the v outer branches of $g = e^2$, with the following properties:*

(1) *Each critical point of f, not a critical point on the infinite branch*

of $f=0$ *is enclosed by some one of the* $u+v$ *surfaces* S_h $(h=1, 2, \cdots, u+v)$.

(2) *Each of the surfaces,* $E_1, E_2, \cdots, E_{p+n}$, *is enclosed by some one of the surfaces* S_h.

(3) *No two of the* $u+v$ *surfaces* S_h *enclose a common point.*

(4) *At each point of* S_α *the outer normal derivative of* g *with respect to the finite space it bounds is positive, and at each point of* S_β *the outer normal derivative of* g *with respect to the finite space it bounds is negative.*

10. *Critical points of* f. $m=0$. Suppose the boundary surfaces, $E_1, E_2, \cdots, E_{p+n}$, of the region D (§ 4) of positive and negative types with respect to g and D, which are enclosed by the surface S_h $(h=1, 2, \cdots, u+v)$ are respectively the following p_h surfaces:

$$E_{h,1},\ E_{h,2}, \cdots, E_{h,p_h};$$

and the following n_h surfaces:

$$E_{h,p_h+1},\ E_{h,p_h+2}, \cdots, E_{h,p_h+n_h}.$$

Then we have obviously

$$\sum_{h=1}^{u+v} (p_h+n_h)=s.$$

Suppose the first Betti number of S_h is $2q_h$. Let D_h designate the finite closed region bounded by S_h and the p_h+n_h surfaces enclosed by S_h. Obviously a boundary surface of D of positive or negative type with respect to g and D, which is a boundary surface of D_h, is of the same type with respect to g and D_h.

Now, from (2), (3), (4) of Lemma 3, the function g is an admissible function in D_h under the new boundary conditions of § 3. By the method used in Case 1 of § 8, we find that in D_α $(\alpha=1, 2, \cdots, u)$ the respective numbers of critical points of type 1 and type 2 of g are

$$\begin{aligned} M_{1,\alpha} &= n_\alpha - 1 + \sum_{\rho=1}^{p_\alpha} g_{\alpha,\rho} + q_\alpha + b_\alpha, \\ M_{2,\alpha} &= p_\alpha + \sum_{\sigma=p_\alpha+1}^{p_\alpha+n_\alpha} g_{\alpha,\sigma} + b_\alpha, \end{aligned} \tag{10.1}$$

where g_{ij} is the first Betti number of E_{ij}, and where b_α is an integer, positive, negative, or zero, but not less than the larger one of the two numbers:

$$-\sum_{\rho=1}^{p_\alpha} g_{\alpha,\rho} - q_\alpha, \quad -\sum_{\sigma=p_\alpha+1}^{p_\alpha+n_\alpha} g_{\alpha,\sigma}.$$

Similarly, by the method used in Case 2 of § 8, we find that in D_β

$(\beta = u + 1, u + 2, \cdots, u + v)$ the respective numbers of critical points of type 1 and type 2 of g are

$$
\begin{aligned}
M_{1,\beta} &= n_\beta + \sum_{\rho=1}^{p_\beta} g_{\beta,\rho} + b_\beta, \\
M_{2,\beta} &= p_\beta - 1 + \sum_{\sigma=p_\beta+1}^{p_\beta+n_\beta} g_{\beta,\sigma} + q_\beta + b_\beta,
\end{aligned} \tag{10.2}
$$

where b_β is an integer, positive, negative, or zero, but not less than the larger one of the two numbers

$$-\sum_{\rho=1}^{p_\beta} g_{\beta,\rho}, \quad -\sum_{\sigma=p_\beta+1}^{p_\beta+n_\beta} g_{\beta,\sigma} - q_\beta.$$

Summing up (10.1) for $\alpha = 1, 2, \cdots, u$, and (10.2) for $\beta = u + 1, u + 2, \cdots, u + v$, we find that the respective numbers of critical points of type 1 and type 2 of g in all the regions D_h $(h = 1, 2, \cdots, u + v)$ are

$$
\begin{aligned}
M_1 &= n - u + \sum_{i=1}^{p} g_i + \sum_{\alpha=1}^{u} q_\alpha + b, \\
M_2 &= p - v + \sum_{j=p+1}^{p+n} g_j + \sum_{\beta=u+1}^{u+v} q_\beta + b,
\end{aligned} \tag{10.3}
$$

where b is an integer, positive, negative, or zero, but not less the larger one of the two numbers:

$$-\sum_{i=1}^{p} g_i - \sum_{\alpha=1}^{u} q_\alpha, \quad -\sum_{j=p+1}^{p+n} g_j - \sum_{\beta=u+1}^{u+v} q_\beta.$$

From (1) of Lemma 3, we have then the following theorem.

THEOREM 3. *Suppose the hypotheses of Theorem 2 hold, and the total mass of the distribution is equal to zero. Suppose the first Betti numbers of the u outer branches S_α $(\alpha = 1, 2, \cdots, u)$ of $g = -e^2$ and the v outer branches S_β $(\beta = u + 1, u + 2, \cdots, u + v)$ of $g = e^2$ (Lemma 3) are $2q_\alpha$ and $2q_\beta$ respectively.*

Then in the space exterior to the surfaces E_k $(k = 1, 2, \cdots, p + n)$ the potential f has at least

$$p + n - (u + v) + \left| \sum_{i=1}^{p} g_i - \sum_{j=p+1}^{p+n} g_j + \sum_{\alpha=1}^{u} q_\alpha - \sum_{\beta=u+1}^{u+v} q_\beta \right|$$

critical points not on the infinite branch of $f = 0$.

The respective numbers M_1 and M_2 of these critical points of type 1 and type 2 are given by (10.3).

It is interesting to note that, so far as formulas go, (8.1) and (8.2) are particular cases of (10.3).

11. *Static charges on conductors in equilibrium.* The case of static charges on conductors (Kellogg, p. 176) in equilibrium is a special case of simple distribution of matter over closed surfaces. Suppose the surfaces E_k, $(k=1, 2, \cdots, p+n)$, of § 4 are the surfaces of $p+n$ finite conductors, and the surface density of the static charges at a point P of any one of the surfaces is $m(P)$. Let $f_n(P)$ designate the outer normal derivative at P of the potential f due to this distribution with respect to the conductor on which P lies. It is well-known that when the distribution is in equilibrium,

$$f_n(P)=4\pi m(P)$$

(Kellogg, p. 164). Hence *an admissible distribution of static charges in equilibrium* over a finite number of finite conductors, the surfaces of which are regular surfaces of class C^3, may de defined as a distribution in equilibrium, the surface density of which is continuous on each conductor and never vanishes. Hence we have the following theorem at once.

THEOREM 4. *The preceding two theorems hold for the potential due to a non-degenerate distribution of static charges in equilibrium over the $p+n$ finite conductors, the surfaces of which are the surfaces, $E_1, E_2, \cdots, E_{p+n}$, of those theorems.*

III. THE NON-DEGENERATE NEWTONIAN POTENTIAL DUE TO A FINITE NUMBER OF POINT MASSES.

12. *The potential. The equipotential surfaces in the neighborhoods of the point masses.* Suppose in the extended 3-space there are p masses $m_i>0$ $(i=1, 2, \cdots, p)$ at the p finite distinct points $P_i=(x^i)$, and n masses $m_j<0$ $(j=p+1, p+2, \cdots, p+n)$ at the n finite points $P_j=(x^j)$, distinct among themselves and from the p points P_i. Then the Newtonian potential due to these point masses is the following function:

$$F(x)=m_k/r_k, \quad (k=1, 2, \cdots, p+n),$$

where r_k is the distance from the point P_k to the variable point (x). The function F becomes positively infinite at P_i and negatively infinite at p_j. The distribution will be said to be *non-degenerate* if F has no degenerate critical point in the finite part of the 3-space.

We shall prove the following lemma.

LEMMA 4. *Suppose F is the potential due to a distribution of p positive point masses at P_i $(i=1, 2, \cdots, p)$ and n negative point masses at P_j $(j=p+1, p+2, \cdots, p+n)$.*

(1) *For a sufficiently large positive constant L, the equipotential locus*

$F(x)=L$ consists of p analytic sphere-like surfaces E_i, E_i enclosing P_i; and the equipotential locus $F(x)=-L$ consists of n analytic sphere-like surfaces E_j, E_j enclosing P_j.

(2) *The finite closed region bounded by each of these sphere-like surfaces is a 3-cell.*

(3) *The function F has no critical point in the finite closed region bounded by any one of these sphere-like surfaces.*

(4) *At each point of E_i [E_j], the outer normal derivative of F with respect to the infinite region bounded by E_i [E_j] is positive [negative].*

Proof. Let us confine our attention on a neighborhood of one of the points P_i. Without loss of generality, we may take one such point at the origin. Then, in a neighborhood of the origin, the potential can be represented in the from:

$$F(x)=u\cdot[1/r-w(x)]$$

where u is a positive constant, the quantity of mass at the origin, r the distance from the origin to the point (x), and $w(x)$ a function analytic in that neighborhood.

Let us introduce the transformation

$$x'_i=x_i/[1-rw(x)],\quad (i=1,2,3). \tag{12.1}$$

The transformation is continuous at the origin, and analytic at any point in a sufficiently small neighborhood of the origin except possibly at the origin. It can be easily shown that the transformations is of class C^1 at the origin, and has a non-vanishing jacobian. Hence the transformation is continuous and one-to-one in a neighborhood N of the origin.

If the positive constant L is taken sufficiently large, the set of points satisfying the inequality $F\geqq L$ has a subset S of points in N and has no point on the boundary of N. Moreover, by virtue of (12.1), we have

$$u\cdot(1/r-w(x))\equiv u/r',$$

where r' is the distance from the transformed origin to (x'). The set S is therefore carried by the transformation (12.1) into the set of points satisfying

$$r'\leqq u/L,$$

which constitutes a 3-cell. Therefore the set S constitutes a 3-cell. The boundary of the set S is, of course, an analytic surface.

It is easy to see that the locus $F(x)=L$ consists of exactly p analytic sphere-like surfaces, one enclosing each positive point mass. Similar statements can be made for $F(x)=-L$.

Thus, the statements (1), (2), and (3) of the lemma follow from the transformation (12.1). Since F has no critical point on these $p+n$ sphere-like surfaces, the normal derivative of F can never vanish on them. Statement (4) then follows at once.

13. *Critical points of F.* The function F and its partial derivatives of the first order are regular at infinity. Thus the Corollary to Theorem 1 holds for F. In fact, the discussions in § 6 and § 7 hold for F almost word for word. So does the discussion in § 9, if we omit the redefinition there because it is unnecessary for F, and if we regard g there as F itself. Hence Lemmas 1-3 with obvious modifications hold for F. On setting g_k $(k=1,2,\cdots,p+n)$ in Theorem 2 equal to zero, we obtain the following theorem.

THEOREM 5. *Suppose there is a non-degenerate distribution of p positive and n negative point masses, the total mass of which is not zero. Then in the finite space the potential F due to this distribution has at least $p+n-1$ critical points. If F has more than this number of critical points, the excess is even.*

If the total mass is negative [positive], F has at least $n-1$ [n] critical points of type one and p [$p-1$] critical points of type two. The excess of critical points of type one, if any, is equal to that of type two.

Similarly, from the proof of Theorem 3 and from the nature of the constants b's there, we have the following theorem.

THEOREM 6. *Suppose there is a non-degenerate distribution of p positive and n negative point masses, the total mass of which is zero. Suppose the first Betti numbers of the u outer branches S_α $(\alpha=1,2,\cdots,u)$ of $F=-e^2$ and the v outer branches S_β $(\beta=u+1,u+2,\cdots,u+v)$ of $F=e^2$ are $2q_\alpha$ and $2q_\beta$ respectively* (cf. Lemma 3).

Then in the finite space the potential F due to this distribution has at least

$$p+n-(u+v)+\sum_{k=1}^{u+v} q_k$$

critical points not on the infinite branch of $F=0$. If F has more than this number of critical points not on the infinite branch of $F=0$, the excess is even.

The respective numbers of these critical points of type 1 and type 2 are

$$M_1=n-u+\sum_{\alpha=1}^{u} q_\alpha+b,\quad M_2=p-v+\sum_{\beta=u+1}^{u+v} q_\beta+b,$$

where b is a positive integer or zero.

COMPLETE SETS OF CONJUGATES UNDER A GROUP.

By G. A. Miller.

1. *Theorems relating to complete sets of conjugate operators.* If a group G is not the identity it cannot involve a complete set of conjugate operators which includes more than one-half of its operators since each operator is transformed into itself by all of its powers. Moreover, when G involves a complete set of conjugate operators which is composed of one-half of its operators then each of them must be of order 2. The product of one of them and any other must be of odd order since this product is transformed into its inverse by each of these two operators of order 2 which are transformed into themselves under G only by their own powers. Hence exactly one-half of the operators of G, including the identity, are of odd order.

The product of two such operators of odd order could not be of order 2 since the products obtained by multiplying one of them successively into all of these operators of order 2 give all of these operators in some order. It therefore results that these operators of odd order constitute a subgroup of index 2 under G. This subgroup must be abelian since each of its operators corresponds to its inverse in an automorphism of G. Moreover, every abelian group of odd order can be extended by an operator of order 2 which transforms each of its operators into its inverse and the resulting group will be such a G. That is, *a necessary and sufficient condition that a group which is not the identity contains a single set of conjugate operators composed of as many as one-half of the operators of the group is that it can be obtained by extending an abelian group of odd order by an operator of order 2 which transforms into its inverse every operator of this abelian group.* This extended group must therefore be either a dihedral group or an extended dihedral group involving a subgroup of odd order and of index 2.

Before determining properties of the groups which involve two complete sets of conjugates such that each set is composed of exactly one-third of the operators of the group it is desirable to establish a theorem which will be very useful in this determination. Suppose that s is an operator of order 3 and that the following equation is satisfied:

$$s^2tst = ts^2ts.$$

From these conditions it follows directly that the three conjugates of t under the powers of s are commutative. It is also clear that the given equation is satisfied whenever each of the two operators st and s^2t is of order 3. Hence

the following theorem: *If s is an operator of order 3 and if t is any operator such that $s^2tst = ts^2ts$ then the three conjugates of t under the powers of s are relatively commutative.*

If G involves two complete sets of conjugate operators such that each of them is composed of one-third of the operators of G then we can select a set of m of these operators, where m is one-third of the order of G, which have the property that the products of a given one of them s_1 into all of them are all of order 3. We shall denote these sets of operators of order 3 by

$$s_1, s_2, \cdots, s_m \qquad s_1^2, s_1s_2, \cdots, s_1s_m$$

respectively. From the theorem noted above it results that these sets have no common operator since it may be assumed without loss of generality that each of these operators is transformed into itself only by its own powers. It also results from this theorem that the order of each of the m distinct operators $1, s_1^2s_2, \cdots, s_1^2s_m$ is not equal to 3. This set of m operators is therefore composed of all the operators of G whose orders are not equal to 3 and the same set would result if we would start with any one of the other operators of order 3 instead of starting with s_1.

Since each of the products obtained by multiplying these m operators successively by one of the remaining operators of G, or by its square, is of order 3, it results that the product of any two of these m operators, or the square of any one of them, is not of order 3. Hence these m operators constitute an invariant subgroup H of index 3 under G whose order is prime to 3. If H involved more than one Sylow subgroup of a given order it would involve two non-commutative operators which would be conjugate under an operator of order 3 of G. This is impossible according to the theorem noted above. Hence we have proved the following theorem: *If a group involves two complete sets of conjugates such that each set is composed of exactly one-third of the operators of the group then it contains an invariant subgroup of index 3 which is the direct product of its Sylow subgroups and the order of each of these subgroups is of the form $3k+1$.*

If a group G of order g involves a complete set of g/n, $n < g$, conjugate operators then the order of its central must divide n and be less than n. In particular, when n is a prime number the central of G is the identity. In this case the order of G cannot be divisible by the square of this prime number since every subgroup of order p^m, p being a prime number, contained in G is found in at least one of the Sylow subgroups of G whose order is a power of p. These statements are illustrated by the theorems noted above and may be employed in giving alternate proofs of these theorems. In particular, it

follows directly therefrom that the order of a group involving a set of conjugates composed of exactly one-half of its operators is divisible by 2 but not by 4 and hence it contains a subgroup of index 2 composed of its operators of odd order including the identity.

2. *Groups involving a small number of complete sets of conjugate non-invariant operators.* There is no group which involves one and only one complete set of conjugate non-invariant operators since such a set could not involve more than one-half of the operators of the group and the central quotient group of a non-abelian group cannot be cyclic. If a group involves just two such sets one of them must involve one-half of the operators of the group since the central could not involve more than one-fourth of these operators. Hence the order of the group is 6 and it is the non-cyclic group of this order. It results directly from the theorem noted above that the number of different complete sets of conjugate non-invariant operators in a group involving a set of conjugates composed of half its operators is $g/4 + \frac{1}{2}$. When the number of these sets of conjugates is a given number $h > 1$ the number of these groups is equal to the number of the abelian groups of order $2h - 1$. In particular, there is at least one such group for every value of $h > 1$.

There is no upper limit for the possible orders of the groups which involve no set of conjugate non-invariant operators since every abelian group has this property. On the contrary, when the number of these sets of conjugates is any given number $k > 1$, the order of the possible groups has an upper limit and hence the number of these groups is also limited. To prove this fact it may be noted that if $g_1, g_2, \cdots, g_k$ represent the orders of the subgroups which transform into itself one of each set of such conjugate operators then the order of the central of G must divide the highest common factor of these g's and it must be less than the smallest of them. In particular, when at least one of these g's is a prime number the central of G must be the identity. As the sum of the reciprocals of these g's is at least equal to $\frac{3}{4}$ at least one of them cannot exceed $4k/3$ and the order of the central can therefore not exceed $2k/3$. After this smallest g has been selected we can proceed similarly with the remaining $k - 1$ g's. Hence the following theorem: *Only a finite number of groups contain any given number of complete sets of conjugate non-invariant operators.*

If a group contains three complete sets of conjugate non-invariant operators and one of its g's is 2 it must be the dihedral group of order 10 according to the theorem noted above. If its smallest g is 3 it is the tetrahedral group, and if its smallest g is 4 it must be a non-abelian group of

order 8. There are therefore four and only four groups which have the property that each of them involves exactly three complete sets of conjugate non-invariant operators. These are the two non-abelian groups of order 8, the tetrahedral group, and the dehedral group of order 10.

If a group contains four complete sets of conjugate non-invariant operators and one of its g's is 2, it must be the dihedral group of order 14. When two of its g's are 3, it is the semi-metacyclic group of order 21, and when only one of them is 3 and another is 4, it may be either the octahedral group or the icosahedral group. It remains to consider the cases when the smallest g exceeds 3. If three of these g's are 4 then G is the holomorph of the group of order 5. When two of them are 4 and the others exceed 4 then G is either the dihedral or the dicyclic group of order 12. It is easy to verify that these are the only possible cases which correspond to a possible group and hence there are exactly seven groups which separately have the property that there appear therein four and only four complete sets of conjugate non-invariant operators.

3. *Groups involving only one complete set of conjugate non-invariant subgroups.* Let G represent any group which involves one and only one complete set of conjugate non-invariant subgroups. The order of these cyclic subgroups must be of the form p^m, p being a prime number. Moreover, there cannot be more than one Sylow subgroup of order q^n, q being different from p, in G and every subgroup of this Sylow subgroup must be invariant thereunder and hence the Sylow subgroup must be abelian when $q > 2$, and either abelian or Hamiltonian when $q = 2$. The subgroup of index p contained in a non-invariant cyclic subgroup of order p^m must be invariant under G and hence a characteristic subgroup of G because it appears in each of its non-invariant subgroups. This subgroup and the Sylow subgroups of G whose orders are prime to p must generate a characteristic subgroup of G, which is the direct product of its Sylow subgroup and involves no invariant operator of G except the identity and possibly operators whose orders are powers of p.

From the preceding paragraph it results directly that the order of G may be divisible by more than two prime numbers and when it is divisible by two such numbers it is of the form p^mq^n, q being a prime of the form $1 + kp$. Moreover, it is obvious that when q is any prime number of this form it is possible to construct one and only one group of order pq^m, where m is an arbitrary natural number, which involves one and only one complete set of conjugate non-invariant subgroups of order p^m. The number of these subgroups is q^n. This proves the following theorem: *A necessary and sufficient condition that a non-abelian group G whose order involves only two*

prime numbers contains one and only one complete set of non-invariant conjugate subgroups is that its order is of the form $p^m q^n$, p and q being prime numbers and q of the form $1+kp$, and that G contains an invariant abelian subgroup of index p and also an operator of order p^m.

It remains to consider the case when the order of G is a power of p. The conjugate cyclic subgroups of order p^m in G must now generate an invariant subgroup H of G while in the preceding case they generate G itself. The former fact results directly from the theorem that in the group generated by a complete set of conjugate subgroups of a prime power group, these subgroups cannot all be conjugate while in the group generated by the Sylow subgroups of a group they are always all conjugate. The quotient group G/H is cyclic. When $p>2$ this follows directly from the theorem that if each of two non-commutative operators whose orders are powers of p transforms the group generated by the other into itself then the product of a power of one of these operators into the other generates a group which is not invariant under the other. This theorem applies also to the case when $p=2$ and at least one of these two operators has an order which exceeds 4. If an operator of G which is not commutative with the subgroup of order 2^m generated by an operator t were of order 4 then t could not be either of order 2 or of order 4. It therefore results that G/H is also cyclic when $p=2$.

The cyclic subgroup of order p^{m-1} which is common to all the conjugate subgroups of order p^m contained in G must be in the central of G since the operators of G which correspond to a generator of G/H must be commutative with all of these operators. The conjugate subgroups of order p^m must therefore correspond to an invariant subgroup of order p in the quotient group corresponding to the group generated by their common subgroup of order p^{m-1}. Hence these conjugate subgroups generate an abelian group and the commutator subgroup of G is of order p. Hence m cannot exceed 1 and H must be the non-cyclic group of order p^2. This proves the following theorem: *If a group of order p^m, p being a prime number, has the property that it contains one and only one complete set of non-invariant conjugate subgroups then it is the non-abelian group which is conformal to the abelian group of this order and of type $(m-1, 1)$.*

It should be noted that each of the two systems of groups which involve together all the groups containing one and only one complete set of non-invariant conjugate subgroups is composed of an infinite number of distinct groups. On the contrary, it was proved above that if a group contains a finite number k of complete sets of non-invariant conjugate operators then its order has an upper limit which is a function of k, and hence the number

of such groups for a given value of k is finite. It may also be noted that it follows directly from the method used above in the study of the groups of order p^mq that there are groups of order p^mq^n which involve exactly n complete sets of non-invariant conjugate subgroups, where n is an arbitrary natural number. To construct such a group it is only necessary to assume that the Sylow subgroup of order q^n is cyclic. This proves the existence of an infinite number of groups which involve separately exactly a given arbitrary number of complete sets of non-invariant conjugate subgroups.

4. *Complete sets of conjugate subgroups of a prime power group.* If in any permutation group of degree n the subgroup composed of all the permutations which omit one letter is of degree $n-k$, $k>1$, then every permutation of this group which omits one letters omits a multiple of k letters. In other words, its systems of imprimitivity composed of k letters each have the property that no permutation of the group permutes some but not all of the letters of a system, and vice versa. If the letters of such a permutation group correspond to a complete set of conjugate non-invariant subgroups or non-invariant operators of a group of order p^m then $k=p^\alpha$, $\alpha>0$. Hence such a complete set of conjugates always involves p^α which have the property that each of them is transformed into itself by every other one of these p^α conjugates. Every operator of the group which transforms one of these p^α conjugates into itself must transform a multiple of p^α of them separately into themselves.

When a group G of order 2^m contains a complete set of 2^{m-2} conjugate subgroups of order 2 the generator s_1 of one of these subgroups must be transformed into itself by the generator s_2 of another and s_1, s_2 generate the four-group. The subgroup of order 8 which contains this four-group must be the octic group since s_1 and s_2 are conjugate thereunder. The subgroup of order 16 which contains this octic group must involve an operator which transforms s_1 into itself multiplied by an operator of order 4 and is commutative with this operator. This transforming operator must therefore be of order 8 and the subgroup of order 16 must be formed by extending the cyclic group of order 8 by an operator of order 2 which transformed its operators either into their seventh or into their third powers. When the order of G exceeds 16 it could not be the latter and the subgroup of order 2 which contains the former, transforms the non-invariant operators of order 2 contained therein into themselves multiplied by an operator of order 8. Hence this group of order 32 contains an operator of order 16. As this process can be continued until G is reached we have established the following theorem: *If a group of order 2^m contains a complete set of 2^{m-2} conjugate subgroups*

of order 2, it is either the dihedral group of order 2^m or the group obtained by extending the cyclic group of order 2^{m-1} by an operator of order 2 which transforms each of its operators into its $2^{m-2}-1$ power. In the former case $m>2$ there are two such sets of conjugates while in the latter case $m>3$ there is only one such set. When $m=3$ the octic group is the only one which satisfies this condition while there are always two such groups when $m>3$.

When a group G of order p^m contains a complete set of p^{m-2} conjugate subgroups of order $p>2$ the generator s_1 of one of these subgroups must again be transformed into itself by the generator s_2 of another and s_1, s_2 generate the non-cyclic group of order p^2. The subgroup of order p^3 which involves this non-cyclic subgroup must be the non-abelian group of this order which involves no operator of order p^2 whenever $m>3$ since s_1 is transformed into itself by only p^2 of the operators of G. This is also true of the subgroup of order p^4 which contains this subgroup of order p^3 when $p>3$ and $m>4$, but when $p=3$ this subgroup of order p^3 involves the abelian group of order p^3 and of type $(2,1)$ while each of its remaining operators is of order 3. It may be noted that when G contains an abelian subgroup of index p then the number of the independent generators of this subgroup cannot exceed p for if this number were larger than p this subgroup would involve a characteristic subgroup of type $(1,1,1,\cdots)$ containing more than p invariants and only one subgroup of order p which would be invariant under G. As this is obviously impossible it results that *when a group of order p^m, p being a prime number, contains a complete set of p^{m-2} conjugate subgroups of order p and also an abelian subgroup of index p then this subgroup cannot involve more than p invariants and when it involves p invariats it must be of type* $(1,1,1,\cdots)$.

When such a subgroup is cyclic and $p>2$ it results directly from the fact that its group of isomorphisms is cyclic that m cannot exceed 3. There is therefore a marked difference between the cases when $p=2$ and when $p>2$ as regards such a cyclic subgroup. In particular, when $p=3$ and this subgroup is cyclic, G is either of order 9 or the non-abelian group of order 27 which involves operators of order 9. When this subgroup is of type $(1,1,1)$ G is the group of order 81 whose only abelian subgroup of order 27 is of this type. On the other hand, when this subgroup has two invariants there is no upper limit to the possible orders of G but the ratio of these two invariants can not exceed 3 in view of the fact that if this ratio would exceed 3, G would involve a characteristic subgroup having for one of its invariants 3 and for the other a number larger than 9. Such a subgroup can obviously not involve an automorphism of order 3 under which only three of its operators are invariant.

TOPOLOGICAL INVARIANCE OF SUB-COMPLEXES OF SINGULARITIES.*

By Arthur B. Brown.

1. *Introduction.* We consider the question of topologically invariant sub-sets of a given complex.† In the first part of the paper we consider sub-complexes which are determined from the incidence relations by simply counting, and here obtain results of which the following is a simple example: If a complex K is an n-cycle (mod. 2), then so is every complex homeomorphic to K. In the second part of the paper, we consider sub-sets defined in simple terms, but such that it is not known at present how to determine them in all cases from the incidence relations. Given two homeomorphic complexes K and K', in terms of these sub-sets we state a necessary and sufficient condition that the image of a given point of K be on a cell of K' of dimension not exceeding i, under all possible homeomorphisms between K and K'.

The proofs depend on the Brouwer theorem of invariance of regionality.

Cells are taken homeomorphic to simplexes, as defined in Veblen's Analysis Situs.‡

2. *Sub-complexes determined by counting.* Given a complex K, let K_n be another name for K, and K_i the sub-complex of K_{i+1} consisting of all the i-cells, and their boundaries, of K_{i+1} that are incident each with a number of $(i+1)$-cells of K_{i+1} not equal to 2. We remark that we might also have K_i include every i-cell, and boundary, of K which is not incident with any cells of K of higher dimensions; and all the succeeding work goes through.

Preparatory to proving the topological invariance of the complexes K_i, we prove two lemmas.

Lemma 1. *Given a complex K and an $(n-1)$-cell a_{n-1} incident with m of its n-cells; then if P is a point on a_{n-1} having a neighborhood on K homeomorphic to an n-cell, m must equal 2.*

Proof. In each of the three other possible cases, $m > 2$, $m = 0$, $m = 1$,

* Presented to the Society, June 13, 1931.

† Analysis situs terminology is as defined in S. Lefschetz's "Topology," *Colloquium Series*, Vol. 12, New York, 1930.

‡ O. Veblen, "Analysis Situs," *The Cambridge Colloquium*, second edition, New York, 1931.

we can easily obtain a contradiction, by use of the theorems of invariance of regionality * and dimensionality.† Therefore m must equal 2.‡

LEMMA 2. *Let K and K' be homeomorphic n-complexes, and b_{n-1} an $(n-1)$-cell of K' which is incident with m n-cells of K', $m \neq 2$. Then under any homeomorphism H between K and K', b_{n-1} is mapped on cells of K of dimensions not exceeding $n-1$; and, neighboring any point P' on it, has images on at least one $(n-1)$-cell of K. Further, if Q' on b_{n-1} corresponds to Q on an $(n-1)$-cell a_{n-1} of K, then a_{n-1} and its incident n-cells correspond to b_{n-1} and its incident n-cells in K', in one-to-one manner, neighboring Q and Q' respectively.*

Proof. Let r be the maximum of the dimensions of the cells of K on which b_{n-1}, neighboring P', is mapped. Then if Q' on b_{n-1} maps on Q of a_r, there must be an $(n-1)$-cell, image of a neighborhood of Q' on b_{n-1}, on a neighborhood of Q on a_r. By the invariance of regionality, r must therefore be at least $n-1$. Since $m \neq 2$, it follows from Lemma 1 that $r \neq n$, hence r must equal $n-1$, as was to be proved.

To prove the final conclusion of the lemma, let Q' on b_{n-1} have as image Q on a_{n-1} of K, and consider the map on K of a neighborhood of Q' on K', say composed of cells one each on b_{n-1} and on the incident n-cells. According to the first part of the lemma, the neighborhood of Q' on b_{n-1} must map on a set neighboring Q on a_{n-1}, hence, by the invariance of regionality, on a neighborhood of Q on a_{n-1}.

Let N'_n be one of the n-cells forming the neighborhood of Q' on K'. The image N_n of N'_n must lie on just one of the n-cells of K incident with a_{n-1}. For, if it had points A and B on two different n-cells, we could join the respective image-points A' and B' by a curve on N'_n and, taking the image curve on K, let C be the first point on that curve, when passing from A to B, which did not lie on the n-cell on which A lay. Then C could not lie on an n-cell, hence would be on a_{n-1}. As the corresponding point C' would lie on N'_n, it would follow from Lemma 1 that a_{n-1} must be incident with just two n-cells of K. Then, considering the homeomorphism of the neighborhoods of Q' and Q, it would follow, again from Lemma 1, that b_{n-1} must be incident with just two

* L. E. J. Brouwer, "Beweis der Invarianz des n-dimensionalen Gebiets," *Mathematischen Annalen*, Vol. 71 (1912), pp. 304-313. Cf. Lefschetz, *loc. cit.*, page 100.

† L. E. J. Brouwer, "Beweis der Invarianz der Dimensionszahl," *Mathematischen Annalen*, Vol. 70 (1911), pp. 161-165. Cf. Lefschetz, *loc. cit.*, page 99. The invariance of dimensionality is of course an immediate corollary of the invariance of regionality.

‡ Cf. proof of invariance of simple circuit, Lefschetz, *loc. cit.*, page 101.

n-cells of K', contrary to hypothesis. Hence N_n can lie on only one of the n-cells of K incident with a_{n-1}.

Further, two different N'_n cannot map on the same n-cell of K, for in that case we could apply the argument just given, with the rôles of K and K' interchanged, and again obtain a contradiction.

Since under H a neighborhood of Q on K must correspond to a neighborhood of Q' on K', we therefore conclude that the final conclusion of Lemma 2 is valid, and the proof is complete.

The following theorem is a special case of Theorem 6 below.

THEOREM 3. *Let K and K' be homeomorphic n-complexes, and M and M' their respective sub-complexes composed of the n-cells and their boundaries. Then under any homeomorphism between K and K', M corresponds to M', and the boundary (mod. 2) of M corresponds to the boundary (mod. 2) of M'.*

Proof. The first conclusion is proved easily by use of the theorem of invariance or regionality. The second conclusion is also proved easily, by use of Lemma 2 and the fact that an $(n-1)$-cell on the boundary of an n-complex is incident with an odd number of n-cells, hence not just 2. We give no further details.

COROLLARY 4. *If K and K' are homeomorphic complexes, then K is an n-cycle (mod. 2) if and only if K' is an n-cycle (mod. 2).**

THEOREM 5. *Let K and K' be two homeomorphic complexes, and K_i, K'_i, $i = 0, 1, \cdots, n$, the respective sub-complexes defined at the beginning of the section. Then, under any homeomorphism H, K_i and K_i' correspond as point sets.*

Proof. Assume the theorem proved for $K_{n-1}, \cdots, K_{i+1}$. The theorem for K_i is then an easy consequence of Lemma 2, taking the n of Lemma 2 to be $i+1$ for the present application. As the case $i = n-1$ requires no assumption, it follows that the theorem is true.

THEOREM 6. *Given a complex K, let A be the sub-complex determined by the i-cells satisfying the following conditions. Each i-cell of A is an i-cell of K_i (defined above), and is incident with a specified number of $(i+1)$-cells of K_{i+1}; each of the latter $(i+1)$-cells is incident with a specified number of $(i+2)$-cells of K_{i+2} (not necessarily the same number in every case); $\cdots$;*

* The corresponding theorem for the n-circuit was proved by Veblen, *loc. cit.*, Chap. III (first or second edition); for the simple n-circuit by Lefschetz, *loc. cit.*, page 101.

and so on, till finally we have $(n-1)$-cells of K_{n-1} incident with specified numbers of n-cells of K. Let K' be a complex homeomorphic to K, and A' the sub-complex of K' defined as described above for A and K, and with the same sets of numbers specified. Then, under any homeomorphism H between K and K', A and A' must correspond.

To prove this theorem, we assume it proved for the sub-complexes similarly defined, but starting with cells of dimension $(i+1)$; and then proceed to prove it for the case stated. The result follows easily upon use of Lemma 2. Hence, as in the preceding proof, the theorem is valid.

3. *Loci of irregular points.* A point P on a complex K will be said to be *p-regular,* or have regularity of order p, if p is the largest integer such that a neighborhood of P on K can be covered by a set of cells all of dimensions at least p; where the set is non-singular on K, and the cells are incident with each other as in a complex. The cells need not be cells of K. The determination of the order of regularity of a given point of K in terms of the incidence relations between the cells of K would involve the solution of some of the outstanding problems of analysis situs; but it is done easily in many cases.

We define J_i, $i=0,1,\cdots,n$, to be the locus of all points on K of regularity i or less.

THEOREM 7. *The locus J_n is all of K, and J_i is a sub-complex of J_{i+1}, of dimension at most i, $i=0,1,\cdots,n-1$.*

The theorem follows easily from the definition, and we omit the proof.

THEOREM 8. *Given K and K' homeomorphic complexes, and J_i and J'_i their respective sub-complexes of points of regularity i or less. Then J_i and J'_i correspond under any homeomorphism between K and K'. Furthermore, a necessary and sufficient condition that a given point P' of K' be mapped on a cell of K of dimension not exceeding i, under all possible homeomorphisms H between K and K', is that P' be on J'_i.*

Proof. The first conclusion is a consequence of the fact that the definition of J_i is purely topological. The sufficiency of the condition in the last conclusion follows from the first part of the theorem. We now prove the necessity.

Let P be the image of P' on K. Since P' is not on J'_i, P is not on J_i, and hence has a neighborhood on K composed of cells of dimensions at least $i+1$ (not necessarily cells of K, of course). Let d_r denote the cell of that neighborhood on which P lies. Since the dimension r of d_r is at least $i+1$,

d_r must contain a point Q as near to P as we like, which does not lie on any cell of K of dimension i or less. This is an easy consequence of the invariance of regionality. It is an easy matter now to define a new homeomorphism H_1 between K and K', which maps P' on Q, but is identical with H except in the neighborhoods of P (including Q) and P'. We need consider H_1 only in these neighborhoods: First we define H_1 over d_r, so as to be identical with H near the boundary of d_r. Next we can define it also over the $(r+1)$-cells of the neighborhood, so as to be continuous over d_r and these $(r+1)$-cells, and identical with H near the (point-set) boundary. Continuing in this way, H_1 is finally determined for the entire neighborhood. Since Q is on an r-cell of K, with $r \geqq i+1$, it follows that the final conclusion of the theorem is valid.

THEOREM 9. *The complex K_i is a sub-complex of J_i.*

Proof. Since K_i and J_i are both sub-complexes of K, and K_i is determined by its i-cells, it is sufficient to prove that if P is a point on an i-cell of K_i, then P lies on J_i. Now if P were not on J_i, we could cover a neighborhood N of P on K by cells of dimensions at least $i+1$, by definition of J_i. Now Theorem 5 holds in the small, as follows from the proofs; hence, relative to N, K_i must coincide with an i-dimensional sub-set of the cells forming N. As the cells forming N would be of dimensions at least $i+1$, K_i could therefore have no points on N, a contradiction to the fact that P on N is on K_i. Hence P must be on J_i, and the theorem is proved.

THEOREM 10. *Theorem 6 holds with K_i replaced by J_i.*

Proof. The same proof applies as in the case of Theorem 6, except for the case of an i-cell of J_i which is incident with zero or two $(i+1)$-cells of J_{i+1}. Here we use the fact that the given i-cell must be mapped on J'_i (Theorem 8), and then Lemma 2 will suffice for completing the proof of the induction. Thus Theorem 10 is valid.

4. *Concluding remarks.* The sub-complexes K_i might easily be made more inclusive. For example, we might let K_i contain all i-cells, and their boundaries, each of which is incident with at least one $(i+1)$-cell of each of the following two classes: (a) $(i+1)$-cells which are not incident with any $(i+2)$-cells; (b) $(i+1)$-cells each of which is incident with at least one $(i+2)$-cells. All the theorems of the paper would remain valid.

Even under an extended definition of K_i, K_i would not in all cases include

all of J_i. Consider, for example, the complex obtained as the join of a point P with a torus.* The point P is not on K_0, K_1 or K_2, but is easily proved to be on J_0, J_1 and J_2.

Finally, we observe that if we consider a complex composed of "cells" each of which, with its boundary, is homeomorphic to an ordinary manifold with regular boundary,† hence not necessarily homeomorphic to a simplex, all the work of the paper goes through unchanged, and all the theorems are valid.

COLUMBIA UNIVERSITY.

* Cf. Veblen, *loc. cit.*, page 96.

† The combinatorial manifold with regular boundary is defined by Lefschetz, *loc. cit.*, page 120. When we limit ourselves to ordinary manifolds, the same definitions apply, except that all cells, spheres, etc., appearing in the definitions of Lefschetz are required to be homeomorphic to ordinary spheres, simplexes, etc. Cf. Veblen, *loc. cit.*, pp. 88 and 92.

A DETERMINATION OF THE TYPES OF PLANAR CREMONA TRANSFORMATIONS WITH NOT MORE THAN 9 *F*-POINTS.

By Mildred E. Taylor.

1. *Introduction.* The types of Cremona transformations can be classified according to the order n or according to the number ρ of the F-points. Coble, [3, p. 83], remarks "that perhaps the latter classification is the more fundamental." Roughly speaking, the range of usefulness of a transformation T is inversely proportional to the number ρ of the F-points. For, unless the F-points of T can be placed at significant points of a given figure, the image is more complicated rather than more simple.

Cremona[6,7] gave a list of transformations for $n = 2$ to $n = 10$. Guccia[8] gave algebraic expressions for the types and their inverses as listed by Cremona. Mlodziejowski [10] listed types for $n = 2$ to $n = 21$. Bianchi [1] by multiplying two De Jonquières transformations obtained the expression for the order n involving 3 parameters. Palatini[12] by multiplying three De Jonquières transformations obtained a new transformation of order n whose expression involves 7 parameters.* He also multiplied four De Jonquières transformations to obtain a new one whose order n involves 15 parameters. In general, if k De Jonquières transformations be compounded the expression for the order n of the new transformation involves $(2^k - 1)$ parameters. Miss Hudson[9] gave types for $n = \alpha\mu - \beta$, where $\beta = 2\gamma + \epsilon$, $\epsilon = 0, 1$. If the number of F-points is 9, there exists an infinite number of transformations. However, in any of the above types when the number of F-points is limited to 9, only a finite number of transformations is obtained. Montesano[11] derived the semi-symmetric types for $\rho = 9$ and obtained independent expressions for each type involving 1 parameter. This paper shows that these types are all related.

In this paper I have obtained explicit algebraic expressions for the integers n, r_i, s_j, α_{ji} attached to every planar Cremona transformation with not more than 9-points, say at $P_9{}^2$, i. e. $p_1, \cdots, p_9$. The method is that used throughout the literature—the composition of known types. These algebraic expressions are classified into *seven* distinct types.

* Miss Hudson [9] states that the expression which Palatini obtained for n by compounding three De Jonquières transformations involves 6 parameters and by compounding four De Jonquières transformations involves 16 parameters.

2. *The Types in the Invariant Subgroup $i_{9,2}$ of the Linear Group $g_{9,2}$.* The product of two Bertini involutions E_2 and E_1 with simple points at p_2, p_1 respectively is called $C_{2,1}$. The semi-symmetric transformation of order 37 with eight 13-fold points and one 4-fold point is called D_9 when the 4-fold point is at p_9. The general element [5, p. 375] of the abelian subgroup of the invariant subgroup $i_{9,2}$ of the linear group $g_{9,2}$ is given by

$$D_2^{\nu_2} D_3^{\nu_3} \cdots D_9^{\nu_9} C_{2,1}^{\rho_2} C_{3,1}^{\rho_3} \cdots C_{9,1}^{\rho_9}, \tag{1}$$

where ν_i is any integer, positive or negative, and $\rho_i = 0,\ 1,\ 2$ with $\Sigma\rho_i \equiv 0 \bmod 3$ $(i = 2, 3, \cdots, 9)$.

By direct multiplication of the factors of (1), explicit algebraic expressions are obtained for n, r_i, s_j, α_{ji} of Type I.

The following notation is used throughout for the elementary symmetric functions of the ν's and the ρ's:

$$\nu = \Sigma_2{}^9 \nu_i, \quad \nu' = \Sigma_2{}^9 \nu_i \nu_j; \quad \rho = \Sigma_2{}^9 \rho_i, \quad \rho' = \Sigma_2{}^9 \rho_i \rho_j; \quad \sigma = \Sigma_2{}^9 \nu_i \rho_i. \tag{2}$$

Also let

$$\gamma = 4\nu^2 + 4\rho^2 - 9\nu' - 4\rho' - 6\sigma. \tag{3}$$

TYPE I.

$$D_2^{\nu_2} D_3^{\nu_3} \cdots D_9^{\nu_9} C_{2,1}^{\rho_2} C_{3,1}^{\rho_3} \cdots C_{9,1}^{\rho_9}.$$

With $i, j = 1, 2, \cdots, 9$ $(i \neq j)$ and $k = 2, 3, \cdots, 9$, let

$$\delta_1 = 2\rho, \quad \delta_k = 3\nu_k - 2\rho_k.$$

Then

$$\begin{aligned} &n = 9\gamma + 1, \quad r_i = 3\gamma + \nu - 3\delta_i, \quad s_j = 3\gamma - \nu + 3\delta_j, \\ &\alpha_{ii} = \gamma - 1, \quad \alpha_{ji} = \gamma + \delta_j - \delta_i. \end{aligned} \tag{4}$$

The remaining elements of $i_{9,2}$ [5, p. 375] are obtained by taking the product of E_1 and (1).

TYPE II.

$$E_1 D_2^{\nu_2} D_3^{\nu_3} \cdots D_9^{\nu_9} C_{2,1}^{\rho_2} C_{3,1}^{\rho_3} \cdots C_{9,1}^{\rho_9}.$$

With i, j, k and δ_1, δ_k as in Type I let

$$\gamma' = \gamma - \nu + 4\rho + 2, \quad \delta'_1 = \delta_1 + 2, \quad \delta'_k = \delta_k.$$

Then

$$(5)\quad \begin{aligned} &n = 9\gamma' + 3\nu - 1, \quad r_i = 3\gamma' + 2\nu - 3\delta_i, \quad s_j = 3\gamma' + 2\nu - 3\delta'_j \\ &\alpha_{ii} = \gamma' + \nu - 2\delta'_i + 1, \quad \alpha_{ji} = \gamma' + \nu - \delta'_i - \delta'_j. \end{aligned}$$

3. *The Remaining Types in the Linear Group* $g_{9,2}$. The factor group $f_{9,2}$ of the group $g_{9,2}$ [5, p. 373] with respect to $i_{9,2}$ is simply isomorphic with the Cremona group $G_{8,2}$ [4, p. 15; 5, p. 349-350]. A study of the transformations in $G_{8,2}$, which is a finite group, shows that each one can be obtained from an element of $i_{9,2}$ by not more than two quadratic transformations. Each type of $i_{9,2}$ is multiplied by a single quadratic, two unrelated quadratic (no F-points in common), or two related quadratic transformations (one F-point in common). There are exactly three different elements of the abelian subgroup of $i_{9,2}$ that will give the same Cremona transformation when multiplied by two unrelated quadratic transformations. Exactly three elements of the non-abelian part of $i_{9,2}$ when multiplied by two unrelated quadratic transformations give the same Cremona transformation. A Cremona transformation which is the product of an element of the abelian subgroup of $i_{9,2}$ and two related quadratic transformations is also the product of some element of the non-abelian part of $i_{9,2}$ and two related quadratic transformations. Hence, there are only *seven* distinct types of planar Cremona transformations.

Since the F-points can be permuted without altering the transformation, the quadratic transformations A_{123}, A_{456}, A_{145} with F-points at p_1, p_2, p_3; p_4, p_5, p_6; p_1, p_4, p_5, respectively are used to simplify the notation.

Type III.

$$D_2^{\nu_2} D_3^{\nu_3} \cdots D_9^{\nu_9} C_{2,1}^{\rho_2} C_{3,1}^{\rho_3} \cdots C_{9,1}^{\rho_9} A_{123}.$$

With $i, j = 1, 2, 3$ $(i \neq j)$ and $k, l = 4, 5, \cdots, 9$ $(k \neq l)$, let

$$\epsilon_{123} = \nu - (\delta_1 + \delta_2 + \delta_3).$$

Then

$$(6)\quad \begin{aligned} &n = 9\gamma - 3\epsilon_{123} + 2, \\ &r_i = 3\gamma + \nu - 3\epsilon_{123} - 3\delta_i + 1, \quad r_k = 3\gamma + \nu - 3\delta_k, \\ &s_j = 3\gamma - \nu - \epsilon_{123} + 3\delta_j + 1, \quad s_l = 3\gamma - \nu - \epsilon_{123} + 3\delta_l, \\ &\alpha_{ii} = \gamma - \epsilon_{123}, \quad \alpha_{kk} = \gamma - 1, \quad \alpha_{ji} = \gamma - \epsilon_{123} + \delta_j - \delta_i + 1, \\ &\alpha_{jk} = \gamma - \epsilon_{123} + \delta_j - \delta_k, \quad \alpha_{li} = \gamma + \delta_l - \delta_i, \quad \alpha_{lk} = \gamma + \delta_l - \delta_k. \end{aligned}$$

Type IV.

$$E_1 D_2^{\nu_2} D_3^{\nu_3} \cdots D_9^{\nu_9} C_{2,1}^{\rho_2} C_{3,1}^{\rho_3} \cdots C_{9,1}^{\rho_9} A_{123}.$$

With γ' and δ''s as in Type II, and i, j, k, l as in Type III, let

$$\epsilon'_{123} = \nu - (\delta'_1 + \delta'_2 + \delta'_3).$$

Then

$$
(7)\quad
\begin{aligned}
&n = 9\gamma' + 3\nu - 3\epsilon'_{123} - 2,\\
&r_i = 3\gamma' + 2\nu - 3\epsilon'_{123} - 3\delta'_i - 1, \quad r_k = 3\gamma' + 2\nu - 3\delta'_k,\\
&s_j = 3\gamma' + 2\nu - \epsilon'_{123} - 3\delta'_j - 1, \quad s_l = 3\gamma' + 2\nu - \epsilon'_{123} - 3\delta'_l,\\
&\alpha_{ii} = \gamma' + \nu - \epsilon'_{123} - 2\delta'_i, \quad \alpha_{kk} = \gamma' + \nu - 2\delta'_k + 1,\\
&\alpha_{ji} = \gamma' + \nu - \epsilon'_{123} - \delta'_i - \delta'_j - 1, \quad \alpha_{li} = \gamma' + \nu - \epsilon'_{123} - \delta'_l - \delta'_i,\\
&\alpha_{jk} = \gamma' + \nu - \delta'_j - \delta'_k, \quad \alpha_{lk} = \gamma' + \nu - \delta'_l - \delta'_k.
\end{aligned}
$$

Type V.

$$D_2^{\nu_2} D_3^{\nu_3} \cdots D_9^{\nu_9} C_{2,1}^{\rho_2} C_{3,1}^{\rho_3} \cdots C_{9,1}^{\rho_9} A_{123} A_{456}.$$

With $i, j = 1, 2, 3$ $(i \neq j)$, $k, l = 4, 5, 6$ $(k \neq l)$, and $m, t = 7, 8, 9$ $(m \neq t)$, let $\epsilon_{456} = \nu - (\delta_4 + \delta_5 + \delta_6)$. Then

$$
(8)\quad
\begin{aligned}
&n = 9\gamma - 6\epsilon_{123} - 3\epsilon_{456} + 4,\\
&r_i = 3\gamma + \nu - 3\epsilon_{123} - 3\delta_i + 1, \quad r_k = 3\gamma + \nu - 3\epsilon_{123} - 3\epsilon_{456} - 3\delta_k + 2,\\
&r_m = 3\gamma + \nu - 3\delta_m, \quad s_j = 3\gamma - \nu - 2\epsilon_{123} - \epsilon_{456} + 3\delta_j + 2,\\
&s_l = 3\gamma - \nu - 2\epsilon_{123} - \epsilon_{456} + 3\delta_l + 1, \quad s_t = 3\gamma - \nu - 2\epsilon_{123} - \epsilon_{456} + 3\delta_t,\\
&\alpha_{ii} = \gamma - \epsilon_{123}, \quad \alpha_{kk} = \gamma - \epsilon_{123} - \epsilon_{456}, \quad \alpha_{mm} = \gamma - 1,\\
&\alpha_{ji} = \gamma - \epsilon_{123} + \delta_j - \delta_i + 1, \quad \alpha_{li} = \gamma - \epsilon_{123} + \delta_k - \delta_i,\\
&\alpha_{ti} = \gamma - \epsilon_{123} + \delta_t - \delta_i, \quad \alpha_{jk} = \gamma - \epsilon_{123} - \epsilon_{456} + \delta_j - \delta_k + 1,\\
&\alpha_{lk} = \gamma - \epsilon_{123} - \epsilon_{456} + \delta_l - \delta_k + 1, \quad \alpha_{tk} = \gamma - \epsilon_{123} - \epsilon_{456} + \delta_t - \delta_k,\\
&\alpha_{jm} = \gamma + \delta_j - \delta_m, \quad \alpha_{lm} = \gamma + \delta_l - \delta_m, \quad \alpha_{tm} = \gamma + \delta_t - \delta_m.
\end{aligned}
$$

Type VI.

$$E_1 D_2^{\nu_2} D_3^{\nu_3} \cdots D_9^{\nu_9} C_{2,1}^{\rho_2} C_{3,1}^{\rho_3} \cdots C_{9,1}^{\rho_9} A_{123} A_{456}.$$

With i, j, k, l, m, t as in Type V and ϵ'_{123} as in Type IV, let

$$\epsilon'_{456} = \nu - (\delta'_4 + \delta'_5 + \delta'_6).$$

Then

$$
(9)\quad
\begin{aligned}
&n = 9\gamma' + 3\nu - 6\epsilon'_{123} - 3\epsilon'_{456} - 4,\\
&r_i = 3\gamma' + 2\nu - 3\epsilon'_{123} - 3\delta'_i - 1, \quad r_k = 3\gamma' + 2\nu - 3\epsilon'_{123} - 3\epsilon'_{456} - 3\delta'_k -\\
&r_m = 3\gamma' + 2\nu - 3\delta'_m, \quad s_j = 3\gamma' + 2\nu - 2\epsilon'_{123} - \epsilon'_{456} - 3\delta'_j + 2,\\
&s_l = 3\gamma' + 2\nu - 2\epsilon'_{123} - \epsilon'_{456} - 3\delta'_l - 1, \quad s_t = 3\gamma' + 2\nu - 2\epsilon'_{123} - \epsilon'_{456} - 3\\
&\alpha_{ii} = \gamma' + \nu - \epsilon'_{123} - 2\delta'_i, \quad \alpha_{kk} = \gamma' + \nu - \epsilon'_{123} - \epsilon'_{456} - 2\delta'_k,\\
&\alpha_{mm} = \gamma' + \nu - 2\delta'_m + 1, \quad \alpha_{ji} = \gamma' + \nu - \epsilon'_{123} - \delta'_i - \delta'_j - 1,\\
&\alpha_{li} = \gamma' + \nu - \epsilon'_{123} - \delta'_l - \delta'_i - 1, \quad \alpha_{ti} = \gamma' + \nu - \epsilon'_{123} - \delta'_t - \delta'_i - 1,
\end{aligned}
$$

$$
\begin{aligned}
&\alpha_{jk} = \gamma' + \nu - \epsilon'_{123} - \epsilon'_{456} - \delta'_j - \delta'_k - 1,\\
&\alpha_{lk} = \gamma' + \nu - \epsilon'_{123} - \epsilon'_{456} - \delta'_l - \delta'_k - 1,\\
&\alpha_{tk} = \gamma' + \nu - \epsilon'_{123} - \epsilon'_{456} - \delta'_t - \delta'_k, \quad \alpha_{jm} = \gamma' + \nu - \delta'_j - \delta'_m,\\
&\alpha_{lm} = \gamma' + \nu - \delta'_l - \delta'_m, \quad \alpha_{tm} = \gamma' + \nu - \delta'_t - \delta'_m.
\end{aligned}
$$

TYPE VII.

$$D_2^{\nu_2} D_3^{\nu_3} \cdots D_9^{\nu_9} C_{2,1}^{\rho_2} C_{3,1}^{\rho_3} \cdots C_{9,1}^{\rho_9} A_{123} A_{145}.$$

With $i, j = 2, 3$ $(i \neq j)$, $k, l = 4, 5$ $(k \neq l)$ and $m, t = 6, 7, 8, 9$ $(m \neq t)$ and ϵ_{123} as in Type III, let $\epsilon_{145} = \nu - (\delta_1 + \delta_4 + \delta_5)$. Then

$$
\begin{aligned}
&n = 9\gamma - 3\epsilon_{123} - 3\epsilon_{145} + 3,\\
&r_1 = 3\gamma + \nu - 3\epsilon_{123} - 3\epsilon_{145} - 3\delta_1 + 2, \quad r_i = 3\gamma + \nu - 3\epsilon_{123} - 3\delta_i + 1,\\
&r_k = 3\gamma + \nu - 3\epsilon_{145} - 3\delta_k + 1, \quad r_m = 3\gamma + \nu - 3\delta_m,\\
&s_1 = 3\gamma - \nu - \epsilon_{123} - \epsilon_{145} + 3\delta_1 + 2, \quad s_j = 3\gamma - \nu - \epsilon_{123} - \epsilon_{145} + 3\delta_j + 1,\\
&s_l = 3\gamma - \nu - \epsilon_{123} - \epsilon_{145} + 3\delta_l + 1, \quad s_t = 3\gamma - \nu - \epsilon_{123} - \epsilon_{145} + 3\delta_t,\\
&\alpha_{11} = \gamma - \epsilon_{123} - \epsilon_{145} + 1, \quad \alpha_{ii} = \gamma - \epsilon_{123}, \quad \alpha_{kk} = \gamma - \epsilon_{145}, \quad \alpha_{mm} = \gamma - 1,\\
(10)\quad &\alpha_{j1} = \gamma - \epsilon_{123} - \epsilon_{145} + \delta_j - \delta_1 + 1, \quad \alpha_{l1} = \gamma - \epsilon_{123} - \epsilon_{145} + \delta_l - \delta_1 + 1,\\
&\alpha_{t1} = \gamma - \epsilon_{123} - \epsilon_{145} + \delta_t - \delta_1, \quad \alpha_{1i} = \gamma - \epsilon_{123} + \delta_1 - \delta_i + 1,\\
&\alpha_{ji} = \gamma - \epsilon_{123} + \delta_j - \delta_i + 1, \quad \alpha_{li} = \gamma - \epsilon_{123} + \delta_l - \delta_i,\\
&\alpha_{ti} = \gamma - \epsilon_{123} + \delta_t - \delta_i, \quad \alpha_{1k} = \gamma - \epsilon_{145} + \delta_1 - \delta_k + 1,\\
&\alpha_{jk} = \gamma - \epsilon_{123} + \delta_j - \delta_k, \quad \alpha_{lk} = \gamma - \epsilon_{145} + \delta_l - \delta_k + 1,\\
&\alpha_{tk} = \gamma - \epsilon_{145} + \delta_t - \delta_k, \quad \alpha_{1m} = \gamma + \delta_1 - \delta_m, \quad \alpha_{jm} = \gamma + \delta_j - \delta_m,\\
&\alpha_{lm} = \gamma + \delta_l - \delta_m, \quad \alpha_{tm} = \gamma + \delta_t - \delta_m.
\end{aligned}
$$

4. *Bertini L-Curves with not more than 9 Multiple Points.* Since the P-curves of the planar Cremona transformations are Bertini L-curves [2; 4, p. 22], the algebraic expressions for r_i, s_j, α_{ji} in (4), (5) $\cdots$ (10) give expressions for the order t and multiplicity t_i of all the Bertini L-curves with not more than 9 multiple points. Each such Bertini curve occurs at least once, but may occur more often in this set of expressions.

BIBLIOGRAPHY.

[1] Bianchi, "Sulle trasformazioni univoche nel piano e nello spazio," *Giornale di Matematiche* (1), Vol. 16 (1878), pp. 263-266.

[2] Bertini, "Sulle curve razionali per le quali si possono assegnare arbitrariamente i punti multipli," *Giornale di Matematiche* (1), Vol. 15 (1877), pp. 329-335.

[3] *Bulletin of the National Research Council No. 63* (1928).

[4] Coble, *American Mathematical Society Colloquium Publications* Vol. 10 (1929).

[5] Coble, "Point Sets and Allied Cremona Groups, Part II," *Transactions of the American Mathematical Society*, Vol. 17 (1916), pp. 345-385.

[6] Cremona, "Sulle trasformazioni geometriche delle figure piane, Nota I," *Giornale di Matematiche*, Vol. I (1863), pp. 305-311.

[7] Cremona, "Sulle trasformazioni geometriche delle figure piane, Nota II," *Giornale di Matematiche*, Vol. 3 (1865), pp. 269-280.

[8] Guccia, "Formole analitiche per la trasformazione Cremoniana," *Rendiconti del Circolo Matematico di Palermo*, Vol. 1 (1884-87), pp. 17-19, 20-23, 24-25, 50-53.

[9] Hudson, "Plane Homaloidal Families of General degree," *Proceedings of the London Mathematical Society* (2), Vol. 22 (1923-1924), pp. 223-243.

[10] Mlodziejowski, "Tables des nombres 21 premiers ordres," *Mathematische Sammlung der Mathematischen Gesellschaft in Moskau*, Vol. 31 (1922), pp. 58-77.

[11] Montesano, "Le corrispondenze birazionali piane emisimmetriche," *Napoli, Reale Accademia delle Scienze Fisiche e Matematiche, Atti* (3), Vol. 24 (1918), pp. 31-78.

[12] Palatini, "Sulle equazioni di condizione delle reti Cremoniane di curve piane," *Periodico di Matematico* (3), Vol. 9 (1912), pp. 129-143.

ON THE REPRESENTATIVE SPACE S_9 OF THE PLANE CUBICS.*

By J. YERUSHALMY.†

In the study of the plane curves of order n it is often convenient to introduce the terminology of projective geometry of hyperspace by representing the curves by points (or hyperplanes) of a space S_N of $N = n(n+3)/2$ dimensions. All the curves satisfying an algebraic condition are then represented by the points of a certain algebraic spread in S_N. Thus all the curves having a double-point are given by a spread of $N-1$ dimensions, those having two double-points by a spread of $N-2$ dimensions, etc.

The representative space S_5 of the conics has been studied extensively by Veronese,‡ Segre,§ Study,¶ and others. In what follows some properties of the S_9 of the plane cubics are found. We obtain the orders of the more important spreads and their multiplicities with respect to each other, and the linear spaces on them with special reference to the surface $F_2{}^9$ representing cubics which degenerate into triple lines. A more detailed study is made of the $F_2{}^9$ by determining the curves on it, its tangent planes and hyperplanes, as well as its double and triple tangent hyperplanes. Most of the results are then extended to the general representative space S_N $[N = n(n+3)/2]$ of the plane n-ics.

We denote by $V_r{}^s$ or $F_r{}^s$ a spread of r dimensions of order s; while a linear space of dimension k is denoted by S_k or R_k.

A $V_8{}^n$ and S_8 in S_9 are called *hypersurface* and *hyperplane* respectively; a $V_2{}^n$, a surface.

For brevity we refer to a point of S_9 as the cubic of the plane which it represents.

1. *The orders of the spreads.* The cubics of a plane π depend, as is known, on one absolute parameter, which is the constant cross-ratio of the 4 tangents which may be drawn to the cubic from a point on it, and which

* Presented at the Chicago Meeting of the American Mathematical Society April 4, 1931.

† National Research Fellow, University of Illinois.

‡ Veronese, "La superficie omaloide, etc.," *Memorie dell' Accademia dei Lincei* (19)3, 1883-4.

§ Segre, "Considerazioni intorno alla geometria della coniche di un piane," *Atti della R. Accademia delle Scienze di Torino*, 20, 1885.

¶ Study, "Über Geometrie der Kegelschnitte," *Mathematische Annalen*, Vol. 27.

is called the *modulus* of the cubic. In fact, if α is the value of this cross-ratio for a cubic, then

$$(1.1) \qquad J = \frac{4(\alpha^2 - \alpha + 1)^3}{(1+\alpha)^2(2-\alpha)^2(1-2\alpha)^2} = \frac{S^3}{T^2}$$

is the only absolute rational invariant for the cubic, where S and T are the two relative invariants of the cubic of degrees 4 and 6 respectively in the coefficients.

That a cubic be of given modulus is therefore equivalent to one condition and we have in S_9 ∞^1 hypersurfaces corresponding to the ∞^1 values of the parameter J. The points of one of these hypersurfaces correspond to all the cubics in π having the same modulus. Corresponding to the values ∞, 0, 1 of J [$\alpha = -1$; $-\epsilon(\epsilon^3 = 1)$; 1] there are three hypersurfaces whose points correspond to harmonic, equianharmonic, and nodal cubics respectively.

We list below the spreads in S_9 in which we will be interested.

a. ∞^1 F_8's of cubics of equal generic modulus.
b. a V_8 of nodal cubics.
c. an F_8 of harmonic cubics.
d. an F_8 of equianharmonic cubics.
e. a V_7 of cubics having 2 double-points (degenerating into a line and a conic).
f. an F_7 of cuspidal cubics.
g. a V_6 of cubics having 3 double-points (composed of 3 lines).
h. an F_6 of cubics having a tacnode (degenerating into a conic and a tangent line).
i. an F_5 of cubics having a triple-point (degenerating into 3 lines of a pencil).
j. an F_4 of cubics composed of a double line and a simple line.
k. an F_2 of cubics degenerating into a triple line.

We recall that the order of a $V_r{}^s$ in S_N is the number, s, of points in which it is met by an arbitrary S_{N-r}. Thus to find the order of any of the above spreads V_k, we look for the number of cubics having the corresponding singularity which are contained in an arbitrary linear system ∞^{9-k} of cubics.

a. Since J (1.1) involves the coefficients of the cubic to the 12-th degree, there are in an arbitrary pencil of cubics 12 cubics of a given generic modulus, and *each of the hypersurfaces is an* $F_8{}^{12}$.

b. In the same way *the hypersurface of nodal cubics is a* $V_8{}^{12}$.

c. The condition that a cubic be harmonic is $T = 0$, which involves the coefficients to the 6-th degree. *This hypersurface is therefore an* $F_8{}^6$.

d. For an equianharmonic cubic $S = 0$, involving the coefficients to the 4-th degree. *This hypersurface is an* F_8^4.

In an arbitrary net of curves of order n there are $\frac{3}{2}(n-1)(n-2) \times (3n^2-3n-11)$ curves having two double-points and $12(n-1)(n-2)$ cuspidal curves.* For $n=3$ we obtain

e. *The spread of cubics degenerating into a conic and a line is a* V_7^{21}.

f. *The spread of cuspidal cubics is an* F_7^{24}.

Caporali † proves that in a linear system ∞^3 of curves of order n, of genus p with σ base-points and of grade N there are

$$\tfrac{1}{6}(N+4p+\sigma)^3-(N+4p+\sigma)(2N+41p+\sigma)$$
$$-\tfrac{1}{6}(175N-3434P+223\sigma)+106$$

curves having 3 double-points and $2(64p-7N-6\sigma+20)$ curves having a tacnode. For $n=3$, $p=1$, $\sigma=0$, $N=9$ we have

g. *The spread of cubics composed of 3 lines of a triangle is a* V_6^{15}.

h. *The spread of cubics degenerating into a conic and a tangent line is an* F_6^{42}.

The order of F_5 of i is the number of cubics with a triple point contained in a generic linear system ∞^4 of cubics

$$\lambda_1\phi_1(x_1x_2x_3)+\lambda_2\phi_2(x_1x_2x_3)+\lambda_3\phi_3(x_1x_2x_3)+\cdots+\lambda_5\phi_5(x_1x_2x_3)=0. \tag{1.2}$$

The condition for a triple-point is that all the derivatives of second order vanish:

$$\sum_{a=1}^{5}\lambda_a\frac{\partial^2\phi_a}{\partial x_1^2}=0,\quad \sum_{a=1}^{5}\lambda_a\frac{\partial^2\phi_a}{\partial x_2^2}=0,\ \cdots \sum_{a=1}^{5}\lambda_a\frac{\partial^2\phi_a}{\partial x_2\partial x_3}=0. \tag{1.3}$$

The number of such curves in the system (1.2) is given by the order of the matrix of 6 rows and 5 columns

---

* Enriques and Chisini, *Lezioni sulla teorie geometria delle equazioni, etc.*, Vol. 22, p. 178.

† Caporali, "Sopra i sistemi lineari triplamante infiniti di curve algebriche piane," *Memorie di Geometria*, Napoli (1888), p. 191-4.

$$(1.4)\qquad \begin{pmatrix} \frac{\partial^2\phi_1}{\partial x_1^2} & \frac{\partial^2\phi_2}{\partial x_1^2} & \cdots & \frac{\partial^2\phi_6}{\partial x_1^2} \\ \frac{\partial^2\phi_2}{\partial x_2^2} & \frac{\partial^2\phi_2}{\partial x_2^2} & \cdots & \frac{\partial^2\phi_6}{\partial x_2^2} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \frac{\partial^2\phi_1}{\partial x_2\partial x_3} & \frac{\partial^2\phi_2}{\partial x_2\partial x_3} & \cdots & \frac{\partial^2\phi_6}{\partial x_2\partial x_3} \end{pmatrix}$$

whose elements are linear in x_1, x_2, x_3. The order of a matrix of n rows and $n+k$ columns in which the order of the element in the i-th row and the j-th column is $l_j+\lambda_i$ is *

$$(1.5)\qquad H_{k+1}=\mu_{k+1}+\mu_k\nu_1+\mu_{k-1}\nu_2+\cdots+\nu_{k+1}$$

where μ_i and ν_i are the elementary symmetric functions of degree i formed from $l_1, l_2, \cdots l_{n+k}$ and $\lambda_1, \lambda_2, \cdots \lambda_{n+k}$ respectively. In our case $l_1=l_2=\cdots=l_6=1$, $\lambda_1=\lambda_2=\cdots=\lambda_6=0$; $H_2=\mu_2=15$. Hence

i. *The spread of cubics with a triple point is an* F_5^{15}.

j1. The order of V_4 of j is the number of cubics composed of a double line and a simple line which are contained in a generic linear system ∞^5 of cubics

$$(1.6)\qquad \lambda_1\phi_1(x_1x_2x_3)+\lambda_2\phi_2(x_1x_2x_3)+\cdots+\lambda_6\phi_6(x_1x_2x_3)=0.$$

This system contains ∞^1 cubics with a triple point. The locus of the triple points is the sextic curve given by

$$(1.7)\qquad C_6\equiv\begin{vmatrix} \frac{\partial^2\phi_1}{\partial x_1^2} & \frac{\partial^2\phi_2}{\partial x_1^2} & \cdots & \frac{\partial^2\phi_6}{\partial x_1^2} \\ \frac{\partial^2\phi_1}{\partial x_2^2} & \frac{\partial^2\phi_2}{\partial x_2^2} & \cdots & \frac{\partial^2\phi_6}{\partial x_2^2} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \frac{\partial^2\phi_1}{\partial x_2\partial x_3} & \frac{\partial^2\phi_2}{\partial x_2\partial x_3} & \cdots & \frac{\partial^2\phi_6}{\partial x_2\partial x_3} \end{vmatrix}=0.$$

This is a general sextic of full genus as can be seen by applying a formula

* This formula was first inferred by Salmon, *Higher Algebra*, and was proved later by F. F. Decker and A. B. Coble. See Coble, "Restricted Systems of Equations," *American Journal of Mathematics*, Vol. 36, p. 410.

given by Coble (*loc. cit.*) which determines the genus of a curve given by a matrix.

Take now an arbitrary line l in the plane π of the system (1.6). Thru any point P on it there passes a linear system ∞^4 of cubics of the system, which contains, by the previous result, 15 cubics having a triple point. Each of these cuts l in 2 other points Q_1 and Q_2. To each point P correspond therefore 30 points Q, and every point Q arises from 30 points P. We have on l a [30, 30] correspondence with 60 coincidences. These are accounted for by (a) the intersections of l with C_6, and (b) by the intersections of l with the double lines of the cubics of the system (1.6) composed of a double line and a simple line.

a) When P is one of the 6 points of intersection of l with C_6, there passes through P one cubic having a triple-point at P. If the point P is considered as belonging to one of the three branches of the cubic, two points Q correspond to it belonging to the other two branches of the curve. But we must consider the point P as belonging successively to the three branches of the curve, and hence such a point absorbs $3 \cdot 2 = 6$ coincidences.

b) When P is an intersection of l with the double-line of a cubic of the system composed of a double-line and a simple line, one point Q coincides with it since, as will be seen later, F_4 is simple on F_5^{15}. The number of such cubics in the system is therefore $60 - 6 \cdot 6 = 24$,* and:

j. *The spread of cubics composed of a double-line and a simple line is an F_4^{24}.*

Corresponding to a linear system ∞^7 of cubics there exists a range of curves of class 3 such that every curve of one system is apolar to every curve of the other system. If a cubic which degenerates into a triple-line is apolar to a curve of class 3 it must be on it. Hence the number of cubics composed of a triple line which are contained in a linear system ∞^7 of cubics is equal to the number of lines common to a range of curves of class 3 which is 9, and

k. *The spread of cubics composed of a line counted three times is an F_2^9.*

---

* Professor Zariski pointed out to me that a more conclusive proof can be given by establishing parametrically the equation of the correspondence. The reasoning given here is the same as the one used by Chasles, *Compte Rendue*, Vol. 64 (1867), p. 799, in establishing the general formula

$$2\mu(n-1) - \nu = rr'(r'-1) + s(s'-1)$$

of which ours is a special case with $\mu = 15$, $\nu = 0$, $r' = 3$, $r = 6$, $s' = 2$. At the end of the paper the order of this spread is obtained by a different method and may be used as a check on this correspondence.

2. *The multiplicities of the spreads on each other.* If a cubic has a singularity α which is the limit form of another singularity β, then the spread of cubics α is on the spread of cubics β and generally with some multiplicity $\mu \geqq 1$. All the hypersurfaces are given by $S^3 - kT^2 = 0$ and for F_7^{24}, $S = T = 0$, it is clear that the F_7^{24}, which is the complete intersection of F_8^6 and F_8^4, is on all the hypersurfaces, and all the spreads of lower dimensions which are on F_7^{24} are therefore also on all the hypersurfaces. The V_8^{12} of nodal cubics contains evidently all the spreads of lower dimensions. It is easily seen that the position of the spreads on one another can be stated thus:

All four of the hypersurfaces contain F_7^{24} but V_8^{12} alone contains V_7^{21}; V_7^{21} contains V_6^{15} and F_6^{42} but F_7^{24} contains only F_6^{42}; and F_6^{42}, V_6^{15}, F_5^{15}, F_4^{24} each contain every spread of lower dimension.

We proceed to determine the multiplicities of most of the spreads on one another. A spread $V_{r_1}^{s_1}$ is of multiplicity μ on a spread $V_{r_2}^{s_2}$ in S_N if an arbitrary S_{N-r_2} on a generic point of $V_{r_1}^{s_1}$ cuts $V_{r_2}^{s_2}$ in $s_2 - \mu$ additional points.

The multiplicities of all the spreads of lower dimension on the hypersurfaces are obtained by direct calculation. We take a pencil of cubics $\phi_1 + \lambda\phi_2 = 0$ and let ϕ_1 take on respectively the form of the various spreads, and substituting in the expression for $S^3 - kT^2 = 0$ we find the multiplicity of the root $\lambda = 0$. Thus we find.

On F_8^4, F_7^{24} and F_6^{42} are simple, F_5^{15} and F_4^{24} are 2-fold, and F_2^9 is 3-fold.

On F_8^6, F_7^{24} is simple, F_6^{42} and F_5^{15} are 2-fold, F_4^{24} is 3-fold, and F_2^9 is 4-fold.

On each F_8^{12}, F_7^{24} is 2-fold, F_6^{42} is 3-fold, F_5^{15} is 4-fold, F_4^{24} is 6-fold, and F_2^9 is 8-fold.

On V_8^{12}, F_7^{24} and V_7^{21} are 2-fold, F_6^{42} and V_6^{15} are 3-fold, F_5^{15} is 4-fold, F_4^{24} is 6-fold, and F_2^9 is 8-fold.

The multiplicity of any spread on F_7^{24} is equal to the product of its multiplicities on F_8^6 and F_8^4 since F_7^{24} is the complete intersection of these two hypersurfaces. Hence

On F_7^{24}, F_6^{42} is 2-fold, F_5^{15} is 4-fold, F_4^{24} is 6-fold, and F_2^9 is 12-fold.

An arbitrary plane α in S_9 cuts out on V_8^{12} a plane curve K of order 12 whose points are in $(1-1)$ correspondence with the points of the Jacobian

J of the net of cubics in π corresponding to α. K is therefore of the same genus as J. If we take α on a point P of any of the spreads which is of multiplicities μ and ν on V_8^{12} and F_7^{24} respectively, K will have a μ-fold point at P and $24-\nu$ double-points on F_7^{24}. If the genus of the corresponding Jacobian J is p, K will have $\gamma = 11 \cdot 10/2 - p - \mu(\mu - 1)/2 - 24 + \nu$ double points on V_7^{21}. $21 - \gamma$ is the multiplicity of the spread on V_7^{21}. Thus we find:

On V_7^{21}, V_6^{15} is 3-fold, F_6^{42} and F_5^{15} are simple.

If in the linear system ∞^4 of cubics (1. 2), ϕ_1 is composed of a double line and a simple line, say $\phi_1 = x_1^2 x_2$ then the elements of the first column in (1. 4) are $2x_2$, 0, 0, $2x_1$, 0, 0 and the order of the matrix is still 15. If, however, ϕ_1 is composed of a triple line, say $\phi_1 = x_1^3$, then the elements of the first column in the above matrix are $6x_2$, 0, 0, 0, 0, 0, and the order of this matrix is equal to the order of the matrix of 5 rows and 4 columns obtained by deleting the first row and column in the above. Using Salmon's formula (1. 5) we find its order to be 10. Hence

On F_5^{15}, F_4^{24} is simple, F_2^9 is 5-fold.

In order to obtain the multiplicity of F_2^9 on F_4^{24}, we take in π a linear system ∞^5 of cubics which contains a cubic degenerating into a triple line a^3, and establish on a line l the same [30, 30] correspondence as in j1. The 60 coincidences in this case are accounted for as follows:

(1) The locus of triple points of the system breaks up into the line a and a quintic C_5. The 5 intersections of l and C_5 give $5 \cdot 6 = 30$ coincidences.

(2) The intersection of l and a absorbs 10 coincidences since F_2^9 is 5-fold on F_5^{15} and a^3 absorbs 5 cubics with a triple point of the system ∞^4 thru (al).

(3) The remaining $60 - 30 - 10 = 20$ coincidences are due to the number of cubics composed of a double line and a simple line in the system. Therefore *F_2^9 is 4-fold on F_4^{24}.*

3. *Linear spaces on the spreads.* To a linear system ∞^k of cubics all of which have the same singularity α corresponds evidently a linear space S_k on the spread of cubics α. The majority of the spreads contain linear spaces. In certain cases they contain one or more rulings, i. e. systems of linear spaces such that through each point of the spread there passes *one* linear space of the system.

Chisini * shows that there exist ∞^8 pencils of cubics of equal generic modulus, ∞^9 pencils of harmonic cubics, and ∞^{11} pencils of equianharmonic cubics in the plane. We have then on the corresponding hypersurfaces ∞^8 lines on each of the ∞^1 F_8^{12}'s, ∞^9 lines on F_8^6, and ∞^{11} lines on F_8^4.

In a recent paper † and in a note ‡ which will be published in a later issue of this Journal, I show that every pencil of equianharmonic cubics is contained in a net of such cubics through 6 of the 9 base-points of the pencil, and that the 9 base-points divide into 3 triples such that through any 2 triples there passes a net of equianharmonic cubics. Hence the ∞^{11} lines of F_8^4 lie on ∞^9 planes in F_8^4 such that there are three planes on every line.

A characteristic property of an equianharmonic cubic is that it is left invariant by a cyclic homology of period 3, the axis of which is a flex-line of the cubic and the center is the point of intersection of the three flex tangents at the three flexes on the axis. For a given homology in the plane there is a linear system ∞^4 of equianharmonic cubics which are invariant for it (if the axis be $x_1=0$ and the center $(1,0,0)$ then the linear system is $x_1^3+\phi_3(x_1x_2)=0$, ϕ_3 being a binary cubic form). Since there are ∞^4 such homologies in the plane there are ∞^4 S_4's on F_8^4. On each point of F_8^4 there are 9 S_4's since each equianharmonic cubic admits 9 homologies into itself.

In determining the linear spaces on the remaining spreads we use two well known theorems of Bertini § which state that if the general curve of a linear system is irreducible, it can have no variable multiple points, and if the general curve is reducible then either all the curves of the system have a fixed part or each is made up of k variable members of a fixed pencil.

The V_8^{12} of nodal cubics is ruled by ∞^2 S_6's corresponding to the linear systems ∞^6 of cubics having a double-point at a given point of the plane. Each S_6 cuts F_2^9 in a cubic curve Γ_3 corresponding to all the cubics composed of the lines on the given point each counted three times. It has ∞^1 planes on F_4^{24}, each plane corresponding to the net of cubics composed of a fixed line on the given point counted twice and an arbitrary line of the plane.

* O. Chisini, "Sui fasci di cubiche a modulo costante," *Rendiconti del Circolo Matematico di Palermo* (1916).

† J. Yerushalmy, "Construction of Pencils of Equianharmonic Cubics," *American Journal of Mathematics*, Vol. 53 (1931), pp. 319-332.

‡ "On the Configuration of the 9 Base-Points of a Pencil of Equianharmonic Cubics."

§ E. Bertini, "Sui sistemi lineari," *Rendiconti dell' Istituto Lombardo* (2), Vol. 15 (1882), pp. 24-28.

It cuts F_5^{15} in an S_8 corresponding to all the cubics having a triple-point at the given point. The cubic Γ_3 of F_2^9 is in this S_3.

The F_7^{24} of cuspidal cubics is ruled by ∞^3 S_4's. Each of these corresponds to a system ∞^4 of cubics having a cusp at a given point of the plane with a fixed cusp tangent. These S_4's are evidently contained in the S_6's of V_8^{12}.

The V_7^{21} of cubics which degenerate into a conic and a line is ruled in two ways. First by ∞^2 S_5's each corresponding to the cubics which degenerate into a fixed line and an arbitrary conic. Every S_5 has a point in common with F_2^9 (the fixed line counted 3 times) and a surface of Veronese (F_2^4) in common with F_4^{24} (corresponding to the cubics composed of the fixed line and an arbitrary double line). There are also on V_7^{21} ∞^5 planes S_2 corresponding to nets of cubics composed of a fixed conic and an arbitrary line of the plane. A generic S_5 meets a generic plane in a point (the point corresponding to the cubic composed of the fixed line of the S_5 and the fixed conic of the S_2) they therefore belong to an S_7 (the ∞^7 of cubics thru the 2 common points of the fixed line and the fixed conic).

The F_6^{42} is ruled by ∞^3 S_3's corresponding to the webs of cubics which degenerate into a fixed line and a web of conics tangent to the line at a given point. Each is contained in an S_5 of V_7^{21}.

On V_6^{15} there are two kinds of planes. There are ∞^4 planes S_2 corresponding to the nets of cubics composed of 2 fixed lines and an arbitrary line of the plane. There are also ∞^4 planes $\bar{S}_2$ corresponding to the nets of cubics composed of a fixed line and two lines varying in a pencil. On each point of V_6^{15} there are 3 planes S_2 and 3 planes $\bar{S}_2$.

The F_5^{15} is ruled by ∞^2 S_3's corresponding to the webs having a triple-point at a given point of the plane. They are contained in the S_6's of V_8^{12} and cut out the cubic curves Γ_3 on F_2^9.

The F_4^{24} is ruled by ∞^2 planes corresponding to the nets of cubics having a fixed double line and a variable line. Each plane has a point in common with F_2^9.

The F_4^{24} contains also ∞^2 surfaces of Veronese (F_2^4) corresponding to the cubics having a fixed line and a variable double line of the plane. Each F_2^4 is contained in an S_5 on V_7^{21}, and cuts F_2^9 in a point.

The F_2^9 contains no linear spaces.

4. *The* F_2^9. We have seen that there are ∞^2 cubics Γ_3 on F_2^9. The points of such a cubic correspond to the cubics composed of a triple-line varying in a pencil. The Γ_3's are ordinary space cubics belonging to the ∞^2 S_3's of F_5^{15}.

The cubics composed of the tangents to a given conic C_2 in π counted 3 times map on the points of a curve Γ_6 of order 6. Such a curve is contained in an S_6 corresponding to the linear system ∞^6 of cubics apolar to the net of curves of class 3 composed of C_2 and an arbitrary point of the plane. It is a rational norm-curve in S_6. There are, evidently, ∞^5 curves Γ_6 on F_2^9.

The ∞^9 hyperplanes S_8 in S_9 cut out on F_2^9 ∞^9 curves Γ_9 of order 9, each an elliptic norm-curve in its S_8.

It is easily verified by referring to π that:

Two cubics Γ_3 intersect in one point. There is a single Γ_3 on any two points of F_2^9 and ∞^1 on one point.

Two curves Γ_6 intersect in 4 points. On any 5 points of F_2^9 there is a single Γ_6 and ∞^1, ∞^2, ∞^3, ∞^4 curves Γ_9 on 4, 3, 2, 1 points respectively.

Two curves Γ_9 intersect in 9 points (the 9 points in which the S_7, common to the two hyperplanes which cut out the two curves, cuts F_2^9). There are 1, ∞^1, $\infty^2, \cdots \infty^8$ curves Γ_9 on respectively 9, 8, 7, $\cdots$ 1 points of F_2^9.

A cubic Γ_3 and a sextic Γ_6 intersect in two points (the 2 points corresponding to the two tangents drawn from the center of the pencil, corresponding to Γ_3, to the conic C_2 counted 3 times).

A cubic Γ_3 and a curve Γ_9 have 3 points in common (the 3 points in which the plane common to S_3 and S_8 cuts Γ_3).

A curve Γ_6 and a curve Γ_9 intersect in 6 points (the 6 points in which the S_5 common to S_6 and S_8 cuts Γ_6).

Take a point P of F_2^9 and the corresponding triple line a^3 in π. A cubic Γ_3 on P corresponds to the lines of a pencil of center Q on a each counted 3 times. To the pencil of cubics composed of the line a counted twice and an arbitrary line of the pencil $\{Q\}$ corresponds a line in S_9 which is tangent to Γ_3 (and also to F_2^9) at P. The ∞^1 tangent lines to the ∞^1 cubics Γ_3 at P are in a plane which is the tangent plane to F_2^9 at P. Hence

The tangent plane T_2 to F_2^9 at a point P (the image of a triple line a^3), corresponds to the net of cubics composed of the line a counted twice and an arbitrary line of the plane.

The ∞^2 tangent planes to F_2^9 generate the F_4^{24}.

An hyperplane S_8 which contains a tangent plane T_2 is a tangent hyperplane to F_2^9 at the point P in which T_2 touches F_2^9. We denote such an hyperplane by R_8. The section of R_8 on F_2^9 is a curve Γ_9 having a double point at the point of contact. It is the generic artional nodal 9-ic curve in R_8.

An hyperplane containing two tangent planes T_2 is a double tangent hyperplane. It contains the S_3 defined by the intersection of the two fixed double lines a_1^3, a_2^3 in π corresponding to the 2 planes T_2^1, T_2^2 (since it contains the 4 linearly independent points a_1^3, a_2^3, $a_1^2a_2$, $a_1a_2^2$) and hence contains the cubic Γ_3 on a_1^3, a_2^3. The section of such an hyperplane degenerates into the cubic Γ_3 and a sextic Γ_6.

An hyperplane containing 3 tangent planes T_2 is a triple tangent hyperplane. It contains the 3 S_3's defined by the 3 points of intersection of the 3 fixed lines in π and its section degenerates into 3 cubics Γ_3.

We proceed to determine in π the linear systems of cubics which map into the various tangent hyperplanes of F_2^9 in S_9.

Consider the linear system ∞^6 of cubics in π passing through 2 points P and Q and tangent there to two lines p and q. On the line PQ take two other points R, S and consider the linear system ∞^5 of cubics having a double point at R and passing simply through S. These two systems have in common the net of cubics composed of the line PQ counted twice and an arbitrary line of the plane. They are therefore contained in a linear system ∞^8 of cubics. The corresponding R_8 in S_9 contains the T_2 corresponding to the common net and is therefore tangent to F_2^9 at the point $\overline{PQ}^3$. There are ∞^8 such hyperplanes in S_9 (we can pick P, Q, R, S, p, q in ∞^8 ways). There are ∞^6 such hyperplanes tangent to F_2^9 at a given point.

Take in π 3 points R, S, T on a line a and another point P not on a. Consider the two linear systems ∞^5 of cubics, the one having R as a double base-point and S as a simple base-point, the other having P and T as double and simple base-points respectively. These two systems have in common the net of cubics which degenerates into the line a and the net of conics on PRS, and belong therefore to a linear system ∞^8 of cubics. The hyperplane R_8 corresponding to this linear system contains the two S_6's on V_8^{12} corresponding to the two linear systems ∞^5 of cubics. It contains therefore the two tangent planes T_2 to F_2^9 at the points a^3 and $\overline{PT}^3$, and is a double tangent hyperplane. Its section degenerates into a cubic Γ_3 and a sextic Γ_6. There are obviously ∞^7 such hyperplanes in S_9, ∞^5 tangent at a given point of F_2^9, and ∞^3 tangent at two given points of F_2^9.

Consider three webs of cubics having a triple point. The corresponding three dimensional spaces $S_3^{(1)}$, $S_3^{(2)}$, $S_3^{(3)}$ on F_5^{15} intersect by twos in three points A, B, C of F_2^9. The $S_6(S_3^{(1)}S_3^{(2)})$ and $S_3^{(3)}$ have the two points B, C in common. They have, hence, the line BC in common and belong to an R_8. This hyperplane contains the three tangent planes T_2 at the points A, B, C and is a triple tangent hyperplane. The 9-ic Γ_9 which it cuts out

on F_2^9 is composed of the three cubics Γ_3 cut out on F_2^9 by $S_3^{(1)}$, $S_3^{(2)}$, $S_3^{(3)}$. There are ∞^6 such R_8 in S_9, ∞^5 tangent to a given point of F_2^9, ∞^2 tangent to any 2 points, and a single R_8 tangent to 3 given points of F_2^9.

If the above 3 webs are such that the three fixed triple points are on a line, then the corresponding spaces $S_3^{(1)}$ $S_3^{(2)}$ $S_3^{(3)}$ meet in a point A of F_2^9. The $S_6(S_3^{(1)}S_3^{(2)})$ contains the tangent lines to $\Gamma_3^{(1)}$ and $\Gamma_3^{(2)}$, it contains therefore the tangent plane T_2 to F_2^9 at A, and since the tangent line to $\Gamma_3^{(3)}$ at A is also in T_2, the three spaces $S_3^{(1)}$ $S_3^{(2)}$ $S_3^{(3)}$ have a line in common and are contained in an R_8. This R_8 is tangent to F_2^9 at A. The 9-ic which it cuts out on F_2^9 has a triple point at A and is composed of the three cubics $\Gamma_3^{(1)}$ $\Gamma_3^{(2)}$ $\Gamma_3^{(3)}$. There are ∞^5 such R_8 in S_9, and ∞^3 tangent at a given point of F_2^9.

The linear system ∞^8 of cubics apolar to a given curve of class 3 degenerating into a double-point P and a simple point Q contains the cubics composed of the lines of the pencils $\{P\}$ and $\{Q\}$ each counted 3 times. It also contains all the cubics which break up into a line of the pencil $\{P\}$ counted twice and an arbitrary line of π. The R_8 corresponding to this system cuts out on F_2^9 two cubics $\Gamma_3^{(1)}$ and $\Gamma_3^{(2)}$ and contains all the tangent planes T_2 at the points of one of these cubics say $\Gamma_3^{(1)}$ and is therefore tangent to F_2^9 along this cubic. Its section with F_2^9 is composed of $\Gamma_3^{(1)}$ counted twice and $\Gamma_3^{(2)}$. There are ∞^4 such R_8 and ∞^2 tangent at any point of F_2^9.

The linear system ∞^8 of cubics passing through a given point P of π contains the cubics which degenerate into the lines of the pencil $\{P\}$ each counted 3 times. It contains also the cubics composed of a line of $\{P\}$ counted twice and an arbitrary ilne of π. The R_8 corresponding to this system cuts out on F_2^9 a cubic Γ_3 and contains the tangent planes T_2 to F_2^9 at all the points of Γ_3. It is therefore a triple tangent hyperplane to F_2^9 along this cubic and its section is made up of Γ_3 counted three times. There are ∞^2 such hyperplanes in S_9, and ∞^1 tangent at any point of F_2^9.

If we consider in the plane π a correlation between its points and its lines, then the system ∞^9 of curves of order 3 in it is transformed into the system ∞^9 of curves of class 3. F_2^9 is then the image of all the curves in π composed of triple points, and there is a $(1-1)$ correspondence between the points of F_2^9 and the points of the plane. If we consider now an hyperplane S_8 which corresponds to a linear system ∞^8 of curves of class 3, as representing the cubic curve (of order 3) apolar to that system, then we have established besides the $(1-1)$ correspondence between the curves of class 3 of π and the points of S_9, also a $(1-1)$ correspondence between the cubic

curves of π and the hyperplanes of S_6. The apolarity condition of a system ∞^k of curves of class 3 and the system ∞^{8-k} of curves of order 3 is expressed by the incidence condition of the corresponding linear spaces. $F_2{}^9$ appears then as the map of the plane π by means of the system ∞^9 of all its cubics. The lines of π go by this mapping into the ∞^2 cubics Γ_3, the conics of π, into the curves Γ_6 and the cubics of π go into the hyperplane sections Γ_9. In general any curve n of π is mapped into a curve of order $3n$. The sections of $F_2{}^9$ by the various tangent hyperplanes R_8 correspond to the cubics having one or more double-points. In general a curve Γ_9 corresponding to a cubic curve with any singularity α is cut out by the hyperplane corresponding to the linear system ∞^8 of curves of class 3 apolar to that cubic.

The $F_2{}^9$ was first met by Del Pezzo * in studying the surfaces of order n immersed in a space of n dimensions. He proves that such a surface is either a ruled surface or $F_2{}^9$ or a projection of it on a space of lower dimensions. $F_2{}^9$ is not a projection of a surface of the same order of a space of higher dimensions. It is a norm-surface. Del Pezzo studies the projection of $F_2{}^9$ on spaces of lower dimensions and in particular its projection in S_8. He obtains the seven surfaces of orders $9, 8, \cdots 3$ depending on the number of points that the center of projection has in common with $F_2{}^9$. If we project $F_2{}^9$ from 6 points on it we obtain in S_3 the general cubic surface. The six centers of projection go into a sextuple of lines on the cubic surface, and since there are no lines on the cubic surface that are skew to the six lines of the sextuple it follows that there are no lines on $F_2{}^9$. Any of the 27 lines of the cubic surface which meets k lines of the sextuple represents a rational curve of order $k+1$ of $F_2{}^9$ which passes through the corresponding k points. The surfaces $\phi^3, \phi^4, \cdots \phi^9$ in S_3 which are the projections of $F_2{}^9$ from r points on it $(r = 6, 5, \cdots 1, 0)$ are the map of the plane by means of a linear system ∞^8 of cubics with r base-points.

Castelnuovo † shows that every non-ruled surface in S_3 whose plane sections are elliptic curves is rational and of order $n \leqq 9$ and is the projection of $F_2{}^9$.

5. *Extesnion to plane curves of order n.* Most of the properties of S_9 can be extended to the S_N $[N = n(n+3)/2]$ whose points represent the plane curves of order n. By the same reasoning as for S_9 we can state that:

---

* P. Del-Pezzo, "Sulle superficie dell' n—*mo* ordine immerse nelle spazio di n dimensioni," *Rendiconti del Circolo Matematico di Palermo*, Vol. 1, p. 241.

† G. Castelnuovo, "Sulle superficie algebriche le cui sezioni piane sono curve ellitiche," *Rendiconti dell' Accademia dei Lincei* (1894), p. 59.

The spread of nodal n-ics is a $V_{N-1}^{3(n-1)^2}$.

The spread of n-ics with 2 double points is a V^{α}_{N-2}

$$\alpha = \tfrac{3}{2}(n-1)(n-2)(3n^2-3n-11).$$

The spread of cuspidal n-ics is an $F_{N-2}^{12(n-1)(n-2)}$.

The spread of n-ics having a tacnode is an F^{β}_{n-3}

$$\beta = 2(25n^2-96n+84).$$

The spread of n-ics with 3 double points is a V^{γ}_{N-3}

$$\gamma = \tfrac{1}{6}\,[(3n^2-6n+4)^3-3(3n^2-6n+4)(45n^2-123n+82) \\ +1542n^2-5151n+4070].$$

The spread of n-ics having a triple-point is an $F_{N-4}^{15(n-2)^2}$, for the elements of the matrix (1.4) are in this case of order $n-2$ and applying Salmon's formula its order is $15(n-2)^2$.

In general *the spread of n-ics having a k-fold point is an* $F^{3\binom{k+2}{4}(n-k+1)^2}_{N+2-[k(k+1)/2]}$.

For the condition for a k-fold point is that the $(k-1)$-th derivatives vanish. The order of the spread is therefore equal to the order of the matrix of $k(k+1)/2$ rows and $k(k+1)/2-1$ columns

$$\begin{pmatrix} \dfrac{\partial^{(k-1)}\phi_1}{\partial x_1^{(k-1)}} & \dfrac{\partial^{(k-1)}\phi_2}{\partial x_1^{(k-1)}} & \cdots & \dfrac{\partial^{(k-1)}\phi_{[(k-1)(k+2)/2]}}{\partial x_1^{(k-1)}} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial^{(k-1)}\phi_1}{\partial x_1^{i}\partial x_2^{j}\partial x_3^{l}} & \cdots & \cdots & \dfrac{\partial^{(k-1)}\phi_{[(k-1)(k+2)/2]}}{\partial x_1^{i}\partial x_2^{j}\partial x_3^{l}} \end{pmatrix}$$

$$(i+j+l=k),$$

whose elements are of order $n-k+1$. Salmon's formula gives for its order the number $3\binom{k+2}{4}(n-k+1)^2$.

For $k=n$ we obtain

The spread of n-ics with an n-fold point is an $F^{3\binom{n+2}{4}}_{n+2}$.

The order of the spread of $n+1$ dimensions of n-ics with an n-fold point composed of a double line and $n-2$ simple lines is obtained by setting up the same correspondence as in the case of the F_4^{24} (j1, § 1). We have in this case on the line l a $\left[3(n-1)\binom{n+2}{4}_1, 3(n-1)\binom{n+2}{4}\right]$ correspondence with $6(n-1)\binom{n+2}{4}$ coincidences. The $n(n+1)/2$ intersection of l with

locus of n-fold points of the system $\infty^{(n-1)(n+2)/2}$ of n-ics given by the determinant

$$\left| \frac{\partial^{(n-1)} \phi_a}{\partial x_1{}^i \partial x_2{}^j \partial x_3{}^l} \right| = 0$$

account for $n(n-1) \cdot n(n+1)/2$ coincidences leaving

$$6(n-1)\binom{n+2}{4} - n^2(n+1)(n-1)/2 = 6(n+1)\binom{n+1}{4}$$

coincidences which arise from the same number of n-ics with an n-fold point, in the system which degenerate into a double line and $n-2$ simple lines. Hence

The spread of n-ics with an n-fold point composed of a double line and $n-2$ simple lines is an $F_{n+1}^{6(n+1)\binom{n+1}{4}}$.

In order to obtain the order of the spread F_4 of 4 dimensions of n-ics which degenerate into an $(n-1)$-fold line and a simple line we take an arbitrary linear space S_{N-4} in S_N. On S_{N-4} there is a web of hyperplanes S_{N-1}. They cut out on $F_2{}^{n^2}$ (the surface of n-fold lines) a web of curves of order n^2 which correspond to a web of curves of order n in the plane. At an intersection of S_{N-4} with the F_4 there is for $n > 2$ one tangent plane T_2 to $F_2{}^{n^2}$. This T_2 and S_{N-4} define an S_{N-2} (since they intersect in a point). On this S_{N-2} there is a pencil of hyperplanes of the web. Each hyperplane of this pencil is a tangent hyperplane to $F_2{}^{n^2}$ at the point of contact of T_2. The number of intersections of S_{N-4} with F_4 (i. e. the order of F_4) is therefore equal to the number of pencils of curves n having a double base-point which is contained in an arbitrary linear system ∞^3 of n-ics

$$\lambda_1\phi_1 + \lambda_2\phi_2 + \cdots + \lambda_4\phi_4 = 0.$$

This number is evidently given by the order of the matrix

$$\begin{pmatrix} \frac{\partial\phi_1}{\partial x_1} & \frac{\partial\phi_2}{\partial x_1} & \cdots & \frac{\partial\phi_4}{\partial x_1} \\ \frac{\partial\phi_1}{\partial x_2} & \frac{\partial\phi_2}{\partial x_2} & \cdots & \frac{\partial\phi_4}{\partial x_2} \\ \frac{\partial\phi_1}{\partial x_3} & \frac{\partial\phi_2}{\partial x_3} & \cdots & \frac{\partial\phi_4}{\partial x_3} \end{pmatrix}$$

of 3 rows and 4 columns whose elements are of order $n-1$. By Salmon's formula this matrix is of order $6(n-1)^2$. Hence

The spread of n-ics composed of an $(n-1)$-fold line and a simple line is an $F_4^{6(n-1)^2}$.

This is a good check on the correspondence by which we obtained the order of F_4^{24}.

The spread of n-ics composed of a line counted n times is an $F_2^{n^2}$.

The points of this surface are in $(1-1)$ correspondence with the point of the plane. It is the map of the plane by means of all the curves of order n. The hyperplane sections are norm-curves of order n and of genus $(n-1)(n-2)/2$. Corresponding to the lines of the plane there are on $F_2^{n^2}$ ∞^2 curves Γ_n of order n. These are rational norm-curves and belong to an S_n on $F_{(n+2)}^{3\binom{n+2}{4}}$. On each point P of $F_2^{n^2}$ there are ∞^1 curves Γ_n. The tangent lines to these at P are in the tangent plane to $F_2^{n^2}$ at this point, which is, in fact, the plane corresponding to the net of n-ics in the plane composed of a fixed line a $(a^3 \equiv P)$ counted $(n-1)$ times and an arbitrary line of the plane. The ∞^2 tangent planes to $F_2^{n^2}$ at all its points generate the $F_4^{6(n-1)^2}$.

The only curves on $F_2^{n^2}$ are those whose order is a multiple of n, (μn), and correspond to the curves of order μ in the plane.

The $F_4^{6(n-1)^2}$ contains ∞^2 surfaces $F_2^{(n-1)^2}$ corresponding to the n-ics in the plane composed of a fixed line a and an arbitrary line counted $n-1$ times. Each such $F_2^{(n-1)^2}$ is contained in an $S_{(n-1)(n+2)/2}$ corresponding to the linear system $\infty^{(n-1)(n+2)/2}$ of n-ics composed of the fixed line a and an arbitrary curve of order $n-1$.

The projection of $F_2^{n^2}$ on S_3 from r points on it $[r=0, 1, 2, \cdots, (n^2+3n-6)/2]$ is a surface of order n^2-r which is the map of the plane by means of a system ∞^3 of curves of order n with r base points.

ON NON-SINGULAR BOUNDED MATRICES.

By Aurel Wintner.

We understand by $A = \| a_{ik} \|$ an infinite matrix which is bounded in the sense of Hilbert and denote the matrix $\| \bar{a}_{ki} \|$ by A^*, so that A^* represents the transposed matrix of the conjugate complex elements. If there exists a bounded matrix B for which both products AB, BA are equal to the unity matrix E then B is, according to Toeplitz, uniquely determined and may therefore be denoted by A^{-1}. We call the bounded matrix A non-singular if the bounded matrix A^{-1} exists. The product of bounded matrices is a bounded matrix. The bounded matrix is Hermitean if $A^* = A$ and unitary if $A^* = A^{-1}$. The unitary matrices form a group of bounded matrices. A bounded matrix is called positive definite if it is Hermitean, non-singular and such that the corresponding Hermitean form cannot take a negative value. The reciprocal matrix of a bounded positive definite matrix always exists and is also positive definite. If V is a unitary and Q a positive definite bounded matrix then VQV^{-1} is also a positive definite bounded matrix.—We shall demonstrate the following theorem:

For any non-singular bounded matrix A there exists exactly one positive definite matrix P and exactly one unitary matrix U for which $A = PU$.

We can call P the modulus and U the argument of A. Since the product of a positive definite bounded matrix and a unitary matrix is always a non-singular bounded matrix, our "polar" representation $A = PU$ yields a one-to-one correspondence between the non-singular bounded matrices A and the *pairs* of bounded positive definite and unitary matrices P, U. We infer that PU is not always equal to UP. Since VPV^{-1} is positive definite and VUV^{-1} unitary for any unitary matrix V it follows from the identity

$$VAV^{-1} = VPUV^{-1} = VPV^{-1}VUV^{-1}$$

and from the uniqueness of the polar components of a non-singular bounded matrix that VPV^{-1} and VUV^{-1} are the polar components of VAV^{-1}. *The unitary invariants of the polar components P, U of A are, therefore, unitary invariants of A itself.*

The demonstration is based on the fact † that every bounded positive definite matrix possesses exactly one bounded positive definite "square root." In other words there exists for any bounded positive definite matrix Q a bounded positive definite matrix P for which $Q = P^2$; furthermore, we have

$$(1) \qquad P_1 = P_2 \quad \text{if} \quad P_1^2 = P_2^2,$$

provided P_1 and P_2 are bounded and positive definite.

Since $(CD)^* = D^*C^*$, we have $(AA^*)^* = AA^*$, i. e. AA^* is always a Hermitean matrix. Furthermore, the Hermitean form belonging to AA^* can be written in the form $\sum_i |\sum_k a_{ik}x_k|^2 \geqq 0$. Finally, AA^* possesses the bounded reciprocal $A^{-1*}A^{-1}$ provided A^{-1} exists. Therefore AA^* is for any non-singular matrix A a positive definite bounded matrix.

Let A be any non-singular bounded matrix. Let P denote the positive definite square-root of AA^*:

$$(2) \qquad P^2 = AA^*.$$

We define a bounded matrix U as follows:

$$(3) \qquad U = P^{-1}A.$$

We then have

$$(3') \qquad U^* = A^*P^{-1},$$

inasmuch as the reciprocal P^{-1} of the Hermitean matrix P is a Hermitean matrix. From (2), (3) and (3′) it follows

$$UU^* = P^{-1}AA^*P^{-1} = P^{-1}P^2P^{-1} = E,$$
$$U^*U = A^*P^{-2}A = A^*(AA^*)^{-1}A = E,$$

i. e. the matrix (3) is a unitary matrix and (3) therefore yields the polar representation $A = PU$.

In order to show that this polar representation of A is unique, suppose that

$$(4) \qquad A = P_1U_1 = P_2U_2.$$

We then have $A^* = U_1^{-1}P_1$, $A^* = U_2^{-1}P_2$ and therefore $AA^* = P_1^2 = P_2^2$, or, by virtue of (1), simply $P_1 = P_2$. Since $P_1 = P_2$ is by supposition positive definite and therefore non-singular equation (4) yields the further condition $U_1 = U_2$ which finishes the demonstration for the uniqueness.

It may be mentioned that the matrix AA^* corresponds to the *dilatation*

† Cf. A. Wintner, "Zur Theorie der beschränkten Bilinearformen," *Mathematische Zeitschrift*, Vol. 30 (1929), p. 267 (below).

ellipsoid of the "deformation" defined by A. The unique resolution of A into the polar components P and U corresponds to the resolution of the *deformation* A into a *dilatation* P and a *rotation* U which are not necessarily commutable. However there also exists a unique representation $A = WR$ where W is a unitary matrix and R a positive definite bounded matrix. To show this we need only write $A = PU = U(U^{-1}PU)$ where $U = W$ is a unitary matrix and $U^{-1}PU = R$ a positive definite bounded matrix. If A is a real matrix the positive definite square root P of AA^* is also real † as is the matrix $P^{-1}A = U$.

It is not difficult to determine the non-singular bounded matrices for which the both polar components P, U are commutable,‡

$$A = PU = UP \qquad (\text{or } P = R;\ U = W). \tag{5}$$

Equation (5) is then and only then fulfilled if $A = PU$ is a so-called normal matrix:

$$AA^* = A^*A. \tag{6}$$

Since (5) can be written in the form

$$A^* = U^{-1}P = PU^{-1} \tag{5$'$}$$

equation (6) follows obviously from (5). Conversely (5) follows from (6). In fact, on introducing $A = PU$, $A^* = U^{-1}P$ in (6) we obtain $P^2 = U^{-1}P^2U$ or $(UPU^{-1})^2 = P^2$, that is, by virtue of (1), simply $UPU^{-1} = P$ which is the same as (5). *The non-singular bounded matrix A is, therefore, then and only then normal if its polar components are commutable.*

In the following we give some additional applications of the positive definite square root.

If $G_1, G_2, \cdots, G_r$ denotes a finite group of bounded matrices then $G_mG_1, G_mG_2, \cdots, G_mG_r$ represents, for any fixed value of m, only a permutation of $G_1, G_2, \cdots, G_r$, so that

$$\sum_{n=1}^{r} G_nG_n^* = \sum_{n=1}^{r} (G_mG_n)(G_mG_n)^* = \sum_{n=1}^{r} G_mG_nG_n^*G_m^*$$

or

$$\Gamma = G_m\Gamma G_m^* \qquad (m = 1, 2, \cdots, r) \tag{7}$$

where

$$\Gamma = \sum_{n=1}^{r} G_nG_n^*.$$

---

† Cf. *loc. cit.*, p. 267.

‡ Cf. *loc. cit.*, p. 282.

The matrix Γ is the sum of positive definite bounded matrices and is, therefore, bounded and positive definite. On denoting by Π the positive definite square root of the positive definite bounded matrix Γ^{-1} we have

$$(8) \qquad \Gamma^{-1} = \Pi^2, \qquad \Pi = \Pi^*.$$

Equations (7) and (8) yield

$$(\Pi G_m \Pi^{-1})(\Pi G_m \Pi^{-1})^* = (\Pi G_m \Pi^{-1})(\Pi^{-1} G_m^* \Pi) = \Pi G_m \Pi^{-2} G_m^* \Pi$$
$$= \Pi G_m \Gamma G_m^* \Pi = \Pi \Gamma \Pi = \Pi \Pi^{-2} \Pi = E$$

and in an analogous manner $(\Pi G_m \Pi^{-1})^*(\Pi G_m \Pi^{-1}) = E$, i. e. $\Pi G_m \Pi^{-1}$ is a unitary matrix. *There, therefore, exists, for any finite group* $G_1, G_2, \cdots, G_r$ *of bounded matrices, a positive definite bounded matrix* Π *in such a manner that the isomorphic group* $\Pi G_1 \Pi^{-1}, \Pi G_2 \Pi^{-1}, \cdots, \Pi G_r \Pi^{-1}$ *contains unitary matrices only.* It may be mentioned that one could apply, just as in the usual demonstration for finite matrices, instead of the positive definite matrix Π, a recursive ("triangle") matrix which is furnished by the Jacobi-Toeplitz reduction of Γ.

If C is a bounded cyclic matrix of the order r (i. e. $E = C^r$ and r the least positive integer p for which $C^p = E$) then

$$C, C^2, \cdots, C^{r-1}, C^r$$

represents a finite group. There exists, therefore, a positive definite bounded matrix Π for which $\Pi C \Pi^{-1}$ is a unitary matrix U. Putting $P = \Pi^{-1}$ we conclude that *there exists, for any bounded matrix* C *for which* $C^r = E$, *a positive definite bounded matrix* P *and a unitary matrix* U *for which* $C = PUP^{-1}$; *furthermore, the spectrum of* C *contains at most* r *numbers and these are always* r*-th roots of unity.* In fact, if $\lambda^r \neq 1$, the formula

$$L = (1 - 1/\lambda^r)^{-1} \sum_{n=0}^{r-1} C^n/\lambda^{n+1}$$

defines a bounded matrix for which we have, by virtue of $C^{n+r} = C^n$, obviously $L(\lambda E - C) = E$ and $(\lambda E - C)L = E$; i. e. $(\lambda E - C)^{-1}$ exists provided $\lambda^r \neq 1$.

If P is a positive definite matrix and U a unitary matrix, the matrix PUP^{-1} is not necessarily a unitary matrix (in accordance with the fact that not every matrix belonging to a finite group is a unitary matrix). In a more precise manner PUP^{-1} is only in the trivial case $PUP^{-1} = U$ a unitary matrix. In fact, if PUP^{-1} is a unitary matrix we have $(PUP^{-1})^* = (PUP^{-1})^{-1}$ or $P^{-1}U^{-1}P = PU^{-1}P^{-1}$, i. e. $P^2 = (UPU^{-1})^2$, and, therefore, according to (1), simply $P = UPU^{-1}$ or $PU = UP$, i. e.

(9) $$PUP^{-1} = U.$$

If T is a non-singular bounded matrix and U_1 a unitary matrix then TU_1T^{-1} is not necessarily a unitary matrix. However, if U_1 and U_2 are unitary matrices for which

(10) $$TU_1T^{-1} = U_2,$$

then there exists a *unitary* matrix W for which

(11) $$WU_1W^{-1} = U_2.$$

The demonstration proceeds as follows. On using the polar representation $T = PW$ of the non-singular bounded matrix T equation (10) takes the form

(10′) $$PWU_1W^{-1}P^{-1} = U_2,$$

i. e. the positive definite bounded matrix P transforms the unitary matrix WU_1W^{-1} in a unitary matrix. From (9) it follows, therefore, $PWU_1W^{-1}P^{-1} = WU_1W^{-1}$, or, according to (10′), simply (11). *Similar unitary matrices are accordingly always unitarily equivalent* or all invariants of the unitary matrices are unitary invariants. The same theorem holds, according to a verbal communication of O. Toeplitz, for equivalent Hermitean matrices (this theorem may be shown also with the use of the polar representation of T). The complete system of the unitary invariants of a bounded Hermitean matrix is due to the Dissertation of Hellinger. It has been shown,† with the use of a trigonometrical momentum problem,‡ that the treatment of the complete system of the unitary invariants of the unitary matrices may be reduced to the Hermitean problem solved by Hellinger.

THE JOHNS HOPKINS UNIVERSITY.

† *Loc. cit.*, p. 274, p. 257.
‡ *Loc. cit.*, p. 269.

CONGRUENT GRAPHS AND THE CONNECTIVITY OF GRAPHS.*

By HASSLER WHITNEY.

Introduction. We give here conditions that two graphs be congruent and some theorems on the connectivity of graphs, and we conclude with some applications to dual graphs. These last theorems might also be proved by topological methods. The definitions and results of a paper by the author on "Non-separable and planar graphs," † will be made use of constantly. We shall refer to this paper as N. For convenience, we shall say two arcs touch if they have a common vertex.

1. *Congruent graphs.* For a definition, see N, 7.‡ In this section we consider no graphs containing 1- or 2-circuits; that is, each arc joins two distinct vertices, and any two vertices are joined by at most a single arc.

THEOREM 1. *Let G and G' be two connected graphs, neither of which consists of three arcs of the form ab, ac, ad. Let there be a 1-1 correspondence between their arcs so that to any two arcs having a common vertex in one graph correspond two arcs having a common vertex in the other. Then G and G' are congruent.*

Case A: *At least one of the graphs, say G, is of one of the following forms:*

$$G_a : \alpha(ab), \beta(ac), \gamma(bc), \delta(bd),$$
$$G_b : \alpha(ab), \beta(ac), \gamma(bc), \delta(bd), \epsilon(cd),$$
$$G_c : \alpha(ab), \beta(ac), \gamma(bc), \delta(bd), \epsilon(cd), \zeta(ad).$$

Consider first the graph G_a. Let G'_a be the corresponding graph, and let $\alpha', \beta', \gamma', \delta'$ be the corresponding arcs of G'_a. As α and β touch, α' and β' touch. Hence they are of the form $\alpha'(a'b')$, $\beta'(a'c')$. δ touches α but not β; hence δ' touches α' but not β'; it has therefore b', but neither a' nor c', as end vertex, and is of the form $\delta'(b'd')$. Now γ touches α, β and δ. Hence γ' touches α', β' and δ'. γ' is not of the form $\gamma'(a'b')$, for then $\alpha'(a'b')$ and

* Presented to the American Mathematical Society, February 28, 1931.

† An outline of this paper will be found in the *Proceedings of the National Academy*, Vol. 17 (1931), pp. 125-127. We refer to this outline as P. The full paper has not yet appeared.

‡ P, 3. We formerly used the term *homeomorphic*.

$\gamma'(a'b')$ would form a 2-circuit. It is thus either $\gamma'(b'c')$ or $\gamma'(a'd')$. In the first case, upon dropping primes, G'_a becomes identical with G_a. In the second case, we interchange the names a' and b', c' and d', β' and δ'. Upon dropping primes, G'_a becomes identical with G_a. Thus in either case, G_a and G'_a are congruent.

Consider next the graph G_b. We form G_b from G_a by adding the arc $\epsilon(cd)$. ϵ touches all the preceding arcs but α. The same must be true of ϵ'. and hence ϵ' is of the form $\epsilon'(c'd')$. Just as above, G_b and G'_b are congruent.

Consider finally the graph G_c. We form it from G_b by adding the arc $\zeta(ad)$. ζ touches all the preceding arcs but γ. The same must be true of ζ'. Hence, if γ' is of the form $\gamma'(b'c')$, ζ' is of the form $\zeta'(a'd')$; if γ' is of the form $\gamma'(a'd')$, ζ' is of the form $\zeta'(b'c')$. In either case, just as before, G_c and G'_c are congruent.

Before proceeding, we shall prove the following

LEMMA. *Let G and G' satisfy the conditions of the theorem. Suppose G contains G_a as a subgraph, and contains also an arc which has just one end vertex in G_a. Let G'_a be the subgraph of G' containing those arcs corresponding to the arcs of G_a. Then G_a and G'_a are congruent, preserving the correspondence between their arcs.*

That is, we can rename the vertices and arcs of G'_a so that arcs that formerly corresponded become identical. By the proof of case A,* G_a and G'_a are congruent, and the correspondence between their arcs is preserved if γ' is of the form $\gamma'(b'c')$. Suppose not. Then γ' is of the form $\gamma'(a'd')$. Let η be an arc of G with one end vertex in G_a. This vertex is either, a, b, c, or d, and η touches α and β; α, γ and δ; β and γ; or δ; but not other arcs of G_a. There is no way of adding an arc to G' (without forming a 2-circuit) so that the similar conditions hold, and we have a contradiction, proving the lemma.

We return now to the theorem.

Case B. *Neither G nor G' is any of the graphs G_a, G_b, G_c.* In this case, we shall prove the stronger theorem, that G and G' are congruent, preserving the correspondence between their arcs. We shall consider first non-separable, then separable, graphs.

Case B1. *At least one of the graphs, say G, is non-separable.*

There are three cases to consider here.

---

* We see, as in Case B1c, that G'_a is connected.

Case B1a. *G is of nullity* 0. By N, theorem 8, G contains but a single arc, ab. Hence G' contains the single arc $a'b'$, and G and G' are congruent.

Case B1b. *G is of nullity* 1. Then, by N, theorem 10,* G is a circuit, and contains therefore at least three arcs. Suppose first G contains just three arcs. They are then of the form $\alpha(ab)$, $\beta(ac)$, $\gamma(bc)$. α' and β' of G' must touch, and they are of the form $\alpha'(a'b')$,$\beta'(a'c')$. γ' must touch both α'. β', and is either of the form $\gamma'(b'c')$ or $\gamma'(a'd')$. But by the hypothesis of the theorem, G' is not of the form $a'b'$, $a'c'$, $a'd'$; it is therefore of the form $a'b'$, $a'c'$, $b'c'$. G and G' are thus congruent, preserving the correspondence between their arcs.

Suppose next G contains at least four arcs, $\alpha(ab)$, $\beta(bc)$, $\gamma(cd)$, $\delta(de)$, $\cdots$; $\zeta(fa)$. Let α', β', γ', δ', $\cdots$, ζ' be the corresponding arcs of G'. α' and β' must touch, and they are therefor $\alpha'(a'b')$, $\beta'(b'c')$. γ touches β but not α; the same must be true of γ', and it is therefore $\gamma'(c'd')$. δ touches γ but neither α nor β; hence δ' is $\delta'(d'e')$. We continue in this manner. Finally, ζ touches the preceding arc and α, but none of the other arcs. The same must be true of ζ', and hence it is of the form $\zeta'(f'a')$. Thus G and G' are congruent as required.

Case B1c. *G is of nullity $N > 1$.* We will assume the theorem is true if G is of nullity $< N$, and will prove it for the case that G is of nullity N, (including the case where $N = 2$). This will establish the proof for this case.

By N, theorem 18,† we can remove an arc or suspended chain C from G, leaving a non-separable graph G_1 of nullity $N - 1$. Let C' be the corresponding arc or arcs of G'. Remove them and any isolated vertices there may then be, forming the graph G'_1. As G_1 is connected, G'_1 is connected. For order the arcs of G_1 so that each arc other than the first touches one of the preceding arcs. Ordering the arcs of G'_1 in the same manner as the corresponding arcs of G_1, each arc other than the first touches one of the preceding arcs; thus G'_1 is connected. C' is a chain. For if C is a single arc, C' is a single arc. Otherwise, the proof is exactly like the proof in the last part of case B1b.

We shall divide case B1c into two cases.

Case B1c$_1$. *None of the vertices of C' other than its two end vertices are in G'_1.* We shall show first that G_1 and G'_1 are congruent, preserving the correspondence between their arcs. G_1 is not of the form ab, ac, ad, as it is non-separable. Suppose G'_1 were of this form. Then each pair of arcs

* P, theorem 6.

† P, theorem 7.

of G_1 have a common vertex, and G_1 is of the form ab, ac, bc. As G contains no 2-circuit, C contains at least two arcs. The first of these, δ say, has just one vertex in G_1, and touches therefor just two of the arcs of G_1. The corresponding arc δ' of C' must touch just two of the arcs $a'b'$, $a'c'$, $a'd'$, and it is therefor of one of the forms $b'c'$, $b'd'$, $c'd'$. But by the hypothesis of this case, δ' has but a single end vertex in G'. This is a contradiction, so G'_1 is not of the form described.

Suppose neither G_1 nor G'_1 are one of the graphs G_a, G_b or G_c of case A. Then the hypotheses of case B1 are fulfilled, and G_1 is of nullity $N-1$, and hence G_1 and G'_1 are congruent as required.

Suppose next one of the graphs G_1, G'_1 was one of the graphs G_a, G_b, G_c. If G_1 is the graph G_c, C contains an arc with just one end vertex in G_c, as G contains no 2-circuit. But G_c contains G_a as a subgraph. Let G'_a be the arcs of G' corresponding to G_a. Then, by the lemma, G_a and G'_a are congruent, preserving the correspondence between their arcs. Hence, following the proof of case A, we see that G_1 and G'_1 are congruent as required. If, next, G_1 is the graph G_b, C contains at least two arcs. For if C contained but one arc, this arc would be $\zeta(ad)$, and G would be the graph G_c. This is ruled out by the hypothesis of case B. Again, G contains G_a as a subgraph, and there is an arc with just one end vertex in G_a. Hence G_a and G'_a, and therefor G_1 and G'_1, are congruent as required. Finally, G_1 is not G_a, for G_a is separable. If now G'_1 were one of the graphs G_a, G_b, G_c, G_1 would be also, by case A, and we are back to the case we have considered. Thus in all cases G_1 and G'_1 are congruent as required.

Rename the vertices and arcs of G'_1 so that upon dropping primes, G'_1 and G_1 become identical. We can do this, preserving the correspondence between the arcs. Suppose the chain C had a and b as end vertices in G_1. We shall show that the end vertices of C' in G'_1 are a' and b'. Let A be the arcs of G_1 on a, and B, those on b. By N, theorem 8, there are at least two arcs in A, as G_1 is of nullity $N-1 \geqq 1$; thus a is the only vertex on them all. Similarly, b is the only vertex on all the arcs B. Let A', B' be the corresponding arcs of G'. They are then on a' and b', respectively.

Suppose first C contained but a single arc, $\alpha(ab)$. As it touches all the arcs of both A and B, the corresponding arc α' of C' touches all the arcs of both A' and B'. Suppose α' were not $\alpha'(a'b')$. Say α' is not on a'. As it must touch all the arcs of A', there are just two such arcs, $a'c'$ and $a'd'$, and α' is $\alpha'(c'd')$. Neither c' nor d' is b'. For then G'_1 would contain the arc $a'b'$, hence G_1 would contain the arc ab, and this arc with the arc α would form a 2-circuit. As $\alpha'(c'd')$ touches all the arcs of B, there are just

two such arcs, $b'c'$ and $b'd'$. There are no more arcs in G'_1. For if there were, as there are no more on a' or b', there must be arcs on c' or d'. Hence there are arcs on c or d in G_1. But then α' would touch these arcs in G'_1, and α would not touch the corresponding arcs in G_1. Hence G' is just the graph G_b. But this is ruled out by the hypotheses of case B. Therefore α' is $\alpha'(a'b')$, and G and G' are congruent as required.

Suppose next C contains at least two arcs. The first arc, α, has a in common with the arcs A. As C' has only its end vertices in G'_1, by hypothesis, the arc α' corresponding to α has just one vertex in G_1, and this vertex must be on all the arcs of A'. It is therefor the vertex a'. Similarly, the other end vertex of C' is b', and thus we can rename the vertices and arcs of C' with primed letters so that, upon dropping primes, it becomes identical with C. Therefore G and G' are congruent, as required. This disposes of case $\mathrm{B1c}_1$.

Case $\mathrm{B1c}_2$. *C' has a vertex besides its two end vertices in G'_1.* Then C' contains just two arcs. For suppose C' contained more than two arcs. Then C would contain more than two arcs. None of these but the first and last touch arcs of G_1. Hence none of the arcs of C'_1 but the first and last have vertices in G'_1, and thus there would be no vertices of C' but the first and last in G'_1.

Let $\alpha(ac)$, $\beta(cb)$ be the two arcs of C, and let $\alpha'(a'c')$, $\beta'(c'b')$ be those of C'. Then a', b' and c' are all in G'_1. There is an arc $\gamma'(c'd')$ in G'_1, where d' is neither a' nor b', as G' contains no 2-circuit. This arc touches both α' and β'. Hence the corresponding arc γ of G touches both α and β, and is therefore $\gamma(ab)$. As G_1 is of nullity $N-1 \geqq 1$, G contains more arcs. As G_1 is non-separable, there must be more arcs on both a and b, for otherwise one of these vertices would be a cut vertex. Let $\delta(ad)$ be an arc on a. It touches α and γ but not β. Hence the corresponding arc δ' of G' is $\delta'(a'd')$. Similarly, if $\epsilon(be)$ is an arc on b in G_1, the corresponding arc in G'_1 is $\epsilon'(b'd')$. As now δ' and ϵ' touch, δ and ϵ touch, and thus e is d. G contains no more arcs. For if it did, it would contain an arc on one of the vertices a, b, d. But we cannot fit a similar arc into G'. G is therefore the graph G_b. But this was ruled out, so case $\mathrm{B1c}_2$ does not exist. Case B1c, and hence case B1, is now disposed of.

Case B2. *Both G and G' are separable.* Let $H_1, \cdots, H_m$ be the components of G, and let $H'_1, \cdots, H'_n$ be those of G'. Let I'_i be the subgraph of G' containing those arcs corresponding to the arcs of H_i, $i = 1, \cdots, m$, and let I_j be the subgraph of G containing those arcs corresponding to the arcs of H'_j, $j = 1, \cdots, n$.

Case B2a. One of the graphs $I_1, \cdots, I_n$; $I'_1, \cdots, I'_m$, say I'_1, is of the form $\alpha'(a'b')$, $\beta'(a'c')$, $\gamma'(a'd')$. Then H_1, being non-separable, is of the form $\alpha(ab)$, $\beta(ac)$, $\gamma(bc)$. There are other arcs in G, and hence there is an arc on one of the vertices a, b, c. Let such an arc be $\delta(ad)$. The corresponding arc δ' in G' is $\delta'(b'c')$. There are no more arcs in G'. For suppose there were an arc with both end vertices in I'_1. It is then either $\epsilon'(b'd')$ or $\zeta'(c'd')$, say the first. The corresponding arc in G must be $\epsilon(bd)$. But then H_1 is not a component of G, for $\alpha(ab)$, $\epsilon(bd)$, $\delta(ad)$ is a circuit with arcs in H_1 and arcs not in H_1, contrary to N, theorem 17.* There cannot be an arc with just one end vertex in I'_1, for the lemma would be contradicted. Thus case B2a does not exist.

Case B2b. *No one of the graphs* $I_1, \cdots, I_n, I'_1, \cdots, I'_m$ *is of the form* ab, ac, ad. Consider first any one of the graphs $H_1, \cdots, H_m$, say H_1, which is one of the graphs G_b, G_c. (No one is G_a, as G_a is separable.) Consider that part of H_1, G_a. As G is connected, there is an arc with just one end vertex in G_a (i. e. in H_1). Hence, by the lemma, G_a and the corresponding arcs of I'_1, G'_a, are congruent, preserving the correspondence between their arcs. Hence, as is seen from the proof of case A, H_1 and I'_1 are congruent similarly. Consider next any component, say H_2, which is not either G_b or G_c. Then the corresponding graph, I'_2, is neither of these graphs, for otherwise, case A would be contradicted. As H_2 is connected, I'_2 is connected. Thus, by case B1, H_2 and I'_2 are congruent, preserving the correspondence between their arcs. Thus H_i and I'_i are congruent as required, $i=1, \cdots, m$. Similarly, H'_j and I_j are congruent as required, $j=1, \cdots, n$.

Now I'_1 is a non-separable subgraph of G', and is therefor contained in one of the components $H'_1, \cdots, H'_n$, say H'_1, of G', by N, theorem 11. Similarly, I_1 is contained in one of the components $H_1, \cdots, H_m$, say H_j, of G. Hence H_1 is contained in H_j, and thus they are identical. Thus the arcs of H_1 and H'_1 correspond, and they are congruent, preserving this correspondence. Similarly for the other components. Rename the vertices and arcs in each H'_i with the subscript i and primes so H'_i becomes identical with H_i when these subscripts and primes are dropped.

Let H_1 and H_2 be two components of G with a common vertex, a, (they have but one common vertex, as otherwise they would form a circuit of graphs, contrary to N, theorem 17 *), and let H_{12} be the graph containing the arcs of both. Let H'_1, H'_2, H'_{12} be the corresponding subgraphs of G'. We shall show that H_{12} and H'_{12} are congruent as required, supposing only that

* P, theorem 5.

H_1, H_2, H'_1, H'_2 are connected. We next show that H_{123}, composed of H_{12} and H_3 together, is congruent as required with the corresponding subgraph H'_{123} of G'. Continuing, we find finally that G and G' are congruent as required.

Let A_1 and A_2 be the arcs of H_1 and H_2 on a. Let A'_1 and A'_2 be the corresponding arcs of H'_1 and H'_2. Then the arcs A'_1 are all on a'_1, and the arcs A'_2 are all on a'_2. We must show that a'_1 and a'_2 are the same vertex in G'. Each arc in A_2 touches an arc in A_1, and hence each arc in A'_2 touches an arc in A'_1. Therefor H'_1 and H'_2 have a common vertex. Call it x'. Moreover, any arc in A'_2 has just one vertex in H'_1, as H'_1 and H'_2 have at most a single vertex in common, being two components of G'.

Suppose first there is but a single arc in H'_1, $a'_1b'_1$. Then if a'_1 is not x', b'_1 is x', and we interchange a'_1 and b'_1 so that a'_1 becomes x'. Suppose next H'_1 contains at least two arcs. If there are two arcs in A'_1, they have only the vertex a'_1 in common, and hence any arc in A'_2 has this vertex as an end vertex; that is, x' is a'_1. If A'_1 contains but a single arc $a'_1b'_1$, there are arcs on b'_1, as H'_1 is connected. ab being the corresponding arc of H_1, there are corresponding arcs on b. The arcs of A_2 touch ab but not these other arcs on b. Hence the arcs of A'_2 touch $a'_1b'_1$ but not these other arcs on b'. They are therefore on the vertex a'_1; that is, x' is a'_1.

Thus in all cases, if x' is not a'_1, we can make it so, preserving the correspondence between the arcs of G and G'. Similarly, if x' is not a'_2, we can make it so. Hence, if a'_1 is not a'_2, we can make it so. As finally, H'_1 and H'_2 have but one common vertex, H_{12} and H'_{12} are congruent as required.

The only fact we need notice to complete the proof is that if I_1 and I_2 are two connected groups of components of G, they have at most a single vertex in common, and the same is true for G'. For otherwise, we could find a circuit entering both of these groups of components, contrary to N, theorem 17.* The proof for case B2b, therefore for case B2, therefore for case B, and therefore of the theorem, is now complete.

THEOREM 2. *Let G and G' be two triply connected graphs, and let there be a 1-1 correspondence between their arcs so that to any set of arcs forming a circuit in one graph, corresponds a set of arcs forming a circuit in the other. Then G and G' are congruent.*

For a definition of "triply connected," see the next section. We shall show that to any two arcs that touch in one graph, there correspond two arcs that touch in the other. Theorem 1 then applies.

* P, theorem 5.

Let $\alpha(ab)$, $\beta(bc)$ be any two arcs in G with the common vertex b. We shall show that the corresponding arcs α', β' in G' have a common vertex. As G is triply connected, we cannot disconnect it without dropping out at least three vertices. Therefore if we drop out the vertex b and the arcs on it, forming the graph G_1, we cannot disconnect this graph without dropping out at least two vertices; that is, G_1 has no cut vertex. Therefore, by N, theorems 5 and 7,* the vertices a and c are contained in a circuit P in G_1. P consists of two chains B and C, joining a and c. Now put back b and the arcs on b. The two arcs ab, bc form a third chain, A, joining a and c. The two chains A and C form a circuit Q, and the two chains A and B form a circuit R. Moreover, no other subset of the arcs of these three chains form a circuit.

The arcs in G' corresponding to these three circuits form three circuits, P', Q', R'. Suppose now the arcs α' and β' had no common vertex. These arcs we shall call A'. Let B' be those arcs corresponding to B, and C', those corresponding to C. Let α', β' be the arcs $\alpha'(a'b')$, $\beta'(c'd')$.

Now as A' and B' together form a circuit, B' consists of two chains, one of them joining a' and either c' or d', say d', and the other joining b' and c'. Similarly, C' is in two chains; either one joins a' and d' and the other joins b' and c', or one joins a' and c' and the other joins b' and d'. The first case cannot be, as then there would be a chain in both B' and C' joining a' and d', and these two chains, or parts of them, form a circuit; but the whole of B' and C' forms a circuit, and therefore no part of them does. Therefore the second case occurs. Consider the following three chains: One part of A', $\alpha'(a'b')$, one part of B', that part joining b' and c', and one part of C', that part joining c' and a'. As B' and C' together form a circuit, they have no vertices in common other than the end vertices of the chains forming them. Hence the three chains described form a circuit, containing arcs of A', B' and C'. The corresponding arcs of G must form a circuit. But this is not the case, as we have seen, and we have a contradiction. Hence α' and β' have a vertex in common, and the theorem is proved.

This theorem is not true for all non-separable graphs. An example is given by the two graphs G', G'', given in N, after the proof of theorem 22.

THEOREM 3. *Let there be a 1-1 correspondence between the arcs of the triply connected graphs G and G' so that to any set of arcs of one graph forming a subgraph of nullity 1 there corresponds a set of arcs of the other graph forming a subgraph of nullity* 1. *Then G and G' are congruent.*

We shall prove that circuits correspond to circuits. Theorem 2 then

* P, theorem 1.

applies. Consider any circuit P of G. It is of nullity 1. Hence the corresponding arcs P' of G' are of nullity 1. Now drop out any arc from P. The nullity is reduced to 0. Hence the corresponding arcs of G' cannot be of nullity 1, and are therefore of nullity 0. Therefore, by N, theorem 9, P' is a circuit. Similarly, if P' is a circuit in G', the corresponding arcs P form a circuit in G.

2. *The connectivity of graphs.* In this section, we allow the graphs to contain 2-circuits, but no 1-circuits.

Definitions. Let G be a graph containing at least $n+1$ vertices, such that it is impossible to drop out $n-1$ or fewer vertices and the arcs on them in such a manner that the resulting graph is not connected. We shall say then that G is n-tuply connected. (We consider only numbers $n \geqq 1$). If G is n-tuply connected but not $(n+1)$-tuply connected, we say its connectivity is n.

For $n=1$, we have: A graph G is simply connected if it contains at least two vertices, and it cannot be disconnected by dropping out no vertices; i. e., G itself is connected. For $n=2$, N, theorems 5 and 6 give: a non-separable graph containing at least three vertices is doubly connected, and conversely.

THEOREM 4. *A necessary and sufficient condition that a graph G be n-tuply connected is that G contain at least $n+1$ vertices, and there exist no two graphs H' and H'' such that for some number $k \leqq n-1$, H' contains the vertices $a'_1, \cdots, a'_k$, b', H'' contains the vertices $a''_1, \cdots, a''_k$, b'', and if we let a'_i and a''_i coalesce, forming the vertex a_i, $i=1, \cdots, k$, we thereby form G.*

For $n=1$, the theorem is: A necessary and sufficient condition that G be simply connected is that G contain at least two vertices, and is not formed of two connected parts H', H''; i. e. G is connected. (For we let $k=n-1=0$ vertices of H' and H'' coalesce to form G.) The theorem is trivial for this case.

Consider now numbers $n>1$. The theorem is trivial if G contains fewer than $n+1$ vertices; assume therefore G contains at least $n+1$ vertices. Suppose first G is formed from two graphs, H', H'' as described. Drop out the vertices $a_1, \cdots, a_k$, and the arcs on them. As H' and H'' had only these vertices in common, b' and b'' are now in different connected parts. Hence G is not n-tuply connected.

Suppose now G is not n-tuply connected. Then we can drop out $k \leqq n-1$ vertices $a_1, \cdots, a_k$, so that there remain two vertices b', b'', which

are in different connected parts in the resulting graph G_1. Let H'_1 be that part of G_1 containing b', and let H''_1 be the rest of G_1. Replace each vertex a_i of G by the two vertices a'_i, a''_i $(i = 1, \cdots, k)$ and replace the arcs we removed in the following manner: An arc $a_i c$, where c is in H'_1, we replace by the arc $a'_i c$. An arc $a_i d$, where d is in H''_1, we replace by the arc $a''_i d$. An arc $a_i a_j$ we replace by the arc $a'_i a'_j$. In this manner we have added the vertices $a'_1, \cdots, a'_k$ to H'_1, forming the graph H', and we have added the vertices $a''_1, \cdots, a''_k$ to H''_1, forming the graph H''. Now H' contains the vertex b' distinct from $a'_1, \cdots, a'_k$, H'' contains the vertex b'' distinct from $a''_1, \cdots, a''_k$, and letting a'_i and a''_i coalesce $(i = 1, \cdots, k)$ forms G. This completes the proof.

THEOREM 5. *If the graph G containing at least two vertices can be disconnected by dropping out $n - 1$ or fewer arcs, it is not n-tuply connected.*

The theorem is trivial if $n = 1$. Assume $n > 1$. Drop out only just enough arcs to disconnect G. The resulting graph is then in two parts, H and H', and each arc we dropped out joins a vertex in H to a vertex in H'. Let these arcs be $a_1 b_1$, and $a_2 b_2, \cdots, a_k b_k$, where $a_1, \cdots, a_k$ are in H and $b_1, \cdots, b_k$ are in H'. Then $k < n$. The vertices $a_1, \cdots, a_k$, and also the vertices $b_1, \cdots, b_k$, may not all be distinct.

Case 1. There are two vertices a_i and b_j which are joined by no arc. Consider first the arc $a_1 b_1$. As it does not join a_i and b_j, either a_1 is distinct from a_i or b_1 is distinct from b_j, say the first. Drop out a_1 and the arcs on it. Consider next the arc $a_2 b_2$. One of the vertices a_2, b_2, is distinct from both a_i and b_j. Drop out this vertex, if it is not already dropped out, and its arcs. Continue in this manner. At the end of the process, we have dropped out fewer than n vertices, and we have dropped out all the arcs $a_1 b_1, \cdots, a_k b_k$. The vertices a_i and b_j are still in the graph, but are joined by no chain. We have thus dropped out less than n vertices and their arcs, disconnecting G, and hence G is not n-tuply connected.

Case 2. Each vertex a_i is joined to each vertex b_j by an arc. Say there are p distinct vertices in the set $a_1, \cdots, a_k$, and q distinct vertices in the set $b_1, \cdots, b_k$. There are then pq arcs in the set $a_1 b_1, \cdots, a_k b_k$, or more, if G contains 2-circuits. Now as $k < n$,

$$pq < n.$$

As

$$pq + 1 - (p + q) = (p - 1)(q - 1) \geqq 0, \ p + q \leqq pq + 1 < n + 1.$$

If G contains n or fewer vertices, it is not n-tuply connected. Assume it

contains at least $n+1$ vertices. Then, as there are but $p+q$ vertices in the set $a_1, \cdots, a_k, b_1, \cdots, b_k$, there is a vertex c in G not in this set.

Dropping out the arcs $a_1b_1, \cdots, a_kb_k$ from G leaves the two connected parts H and H', one of which, say H, contains the vertex c. Drop out then the vertices $a_1, \cdots, a_k$ and their arcs from G. We have dropped out fewer than n vertices and their arcs, disconnecting G, as b_1 and c are joined by no chain, and thus G is not n-tuply connected.

THEOREM 6. *If a vertex and its arcs be dropped out of the n-tuply connected graph G, the resulting graph G' is $(n-1)$-tuply connected.*

For G' contains at least n vertices, and we must drop out at least $n-1$ vertices to disconnect it.

THEOREM 7. *A necessary and sufficient condition that a graph containing no 2-circuit be n-tuply connected is that any two of its vertices be joined by n distinct chains.**

For $n=1$, the theorem is obvious. Suppose $n=2$. If the graph contains but two vertices, the theorem is obvious. If the graph contains at least three vertices, the theorem becomes theorem 7 of N.† Assume the theorem is true for all numbers $1, 2, \cdots, n-1$; we shall prove it for $n=n$. This will establish it in general.

We shall prove first the sufficiency of the condition. To show that G contains at least $n+1$ vertices, let a and b be any two of its vertices; they are joined by n distinct chains. As there is at most a single arc ab, there must be $n-1$ distinct chains joining a and b, each of which contains at least two arcs. Each chain contains at least one vertex in its interior, so there are at least $n-1+2=n+1$ vertices in G. Now suppose we could disconnect G by dropping out $k<n$ vertices. Some two remaining vertices, a, b, are now joined by no chain. Hence in the original graph each chain joining them must pass through one of the k vertices we dropped out, and there are not n distinct chains therefore, a contradiction.

To prove the necessity of the condition, we must show that for $n \geqq 2$, if G is an n-tuply connected graph and a, f are any two of its vertices, then a and f are joined by n distinct chains. As $n>2$, G is non-separable, and

* i. e. chains which have only the two given vertices in common. Similar theorems have been proved by K. Menger, "Zur allgemeinen Kurventheorie," *Fundamenta Mathematicae*, Vol. 10 (1926), pp. 96-115, Satz β; N. E. Rutt, "Concerning the Cut-Points of a Continuous Curve When the Arc Curve ab Contains Exactly n Independent Arcs," *American Journal of Mathematics*, Vol. 51 (1929), pp. 217-246.

† P, theorem 1.

hence, by N, theorem 7,* there is a circuit passing through a and f. One of the two chains joining them, say the chain $D = ab, bc, cd, \cdots, ef$, contains at least two arcs. Form the graph G_2 from G by dropping out the vertex b and the arcs on it. By theorem 6, G_2 is $(n-1)$-tuply connected, and hence there are $n-1$ distinct chains joining a and c. Now replace the vertex b and the arcs on b. The arcs ab, bc form another chain from a to c, so that there are in G n distinct chains joining a and c. Now if f is not joined to a by n distinct chains, there is a first vertex of the chain D (coming after c), say d, which is not, while the vertex preceding it, say c, is. We shall show this leads to a contradiction.

Let $A_1, A_2, \cdots, A_n$ be n distinct chains from a to c. Form the graph G_1 from G by dropping out the vertex c and the arcs on c. There are in G_1 $n-1$ distinct chains $B_1, B_2, \cdots, B_{n-1}$ from d to a. We shall find in G n distinct chains $C_1, C_2, \cdots, C_n$ from d to a.

Case 1. d is not on any of the chains $A_1, \cdots, A_n$. The method is to find first a set of chains $C'_1, \cdots, C'_{n-1}$, with the following properties:

(1) For some number l, $0 \leqq l \leqq n-1$, l of the chains, say C'_{n-l}, $C'_{n-l+1}, \cdots, C'_{n-1}$ are identical with the chains $B_{n-l}, B_{n-l+1}, \cdots, B_{n-1}$.

(2) Each other chain C'_i, (of which there are $n-1-l$), is identical with a part of B_i, that part stretching from d to a certain vertex, say b_i, of one of the chains $A_1, \cdots, A_n$, say A_i, and these vertices lie on distinct chains.

(3) No one of the chains C'_i has a vertex in common with either a vertex of that part of any chain A_j lying between a and b_j, $j = 1, 2, \cdots$, $n-1-l$, or with any vertex of any of the chains $A_{n-l}, A_{n-l+1}, \cdots, A_n$, other than the vertex a. (Of course no chain B_i passes through the vertex c).

We can then construct the chains $C_1, \cdots, C_n$ as follows: For $i = 1$, $2, \cdots, n-1-l$, C_i consists of C'_i plus that part of A_i lying between a and b_i; for $j = n-l, n-l+1, \cdots, n-1$, C_j is exactly C'_j; and C_n consists of the arc cd plus the chain A_n. These are n distinct chains from d to a, as required.

We shall find the chains $C'_1, \cdots, C'_{n-1}$ in the following manner: First follow B_1 from d towards a till we reach a vertex of one of the chains $A_1, \cdots, A_n$, or the vertex a itself. This much of B_1 might serve as C'_1. Next follow B_2 towards a, and alter C'_1 if necessary, so we have two chains that might serve as C'_1, C'_2. Next follow B_3 towards a, and alter the other two chains if necessary till we have three distinct chains which might serve as C'_1, C'_2, C'_3. Continue in this manner till we have found the $n-1$ re-

* P, theorem 1.

quired chains. At each stage therefore, the properties (1), (2), (3) hold, with $n-1$ replaced by some number $k \leqq n-1$.

To turn to the actual construction, we find C'_1 as described. Next follow B_2 from d towards a. Suppose first B_1 was entirely distinct from the chains $A_1, \cdots, A_n$, so that C'_1 is B_1. If we reach a before touching any of the chains $A_1, \cdots, A_n$, B_2 forms C'_2. If we reach first a vertex b_2 of the chain A_2, say, this much of B_2 forms C'_2. In either case, (1), (2) and (3) hold. Suppose next B_1 had the vertex b_1 in common with A_1 say. If we reach either the vertex a or a vertex b_2 of one of the chains $A_2, \cdots, A_n$, say A_2, before reaching any other vertex of the chains $A_2, \cdots, A_n$, or any vertex of $A_1(ab_1)$, (that is, that part of A_1 lying between a and b_1), this much of B_2, that is, B_2 itself or $B_2(db_2)$, forms C'_2, and again (1), (2) and (3) hold. We are thus in difficulty only if we reach a vertex b'_2 of $A_1(ab_1)$ before reaching either a or a vertex of one of the chains $A_2, \cdots, A_n$.

In this case, let C'_2 be $B_2(db'_2)$. Let us follow B_1 further from b_1 towards a. If we reach either a or a vertex b'_1 of one of the chains $A_2, \cdots, A_n$ before reaching a vertex of $A_1(ab'_2)$, this much of B_1 added to the original C'_1 forms the new C'_1, and again (1), (2) and (3) hold. We are in difficulty only if we reach first a vertex, say b''_1, of $A_1(ab'_2)$. In this case, let $B_1(db''_1)$ be the new C'_1. Follow now B_2 from b'_2 towards a, forming more of the new C'_2. Again, we are in difficulty only if we reach a vertex of $A_1(ab''_1)$ before reaching either a or a vertex of one of the chains $A_2, \cdots, A_n$. Continuing, the process must at some time come to an end, as there are only a finite number of vertices in the chain A_1. We now have two chains C'_1, C'_2, as required.

We shall assume now we have found $k-1$ chains $C'_1, \cdots, C'_{k-1}$, as required, and we shall find the k'th chain C'_k. We can rename chains and vertices so that for some number $l \leqq k-1$, l of the chains, C'_{k-l}, C'_{k-l+1}, $\cdots$, C'_{k-1} are identical with B'_{k-l}, $B'_{k-l+1}, \cdots, B'_{k-1}$, respectively, and each other chain C'_i is identical with $B_i(db_i)$, $i=1, 2, \cdots, k-1-l$, where b_i lies on the chain A_i.

Follow B_k from d towards a. If we reach either the vertex a or a vertex b_{k-l} of one of the chains A_{k-l}, $A_{k-l+1}, \cdots, A_n$, say A_{k-l}, before reaching any other vertex of these chains or any vertex of $A_i(ab_i)$, $i=1, 2, \cdots, k-1-l$, this much of B_k forms C'_k, and (1), (2) and (3) hold. We are in difficulty only if we reach first a vertex b' of $A_i(ab_i)$, $1 \leqq i \leqq k-1-l$. Let this much of B_k be C'_k. We alter C'_i as follows: Follow B_i from b_i towards a. If we reach a or a vertex of one of the chains $A_{k-l}, \cdots, A_n$ first, this much of B_i together with $B_i(db_i)$ forms the new C'_i, and again (1), (2) and (3) hold. Suppose we reached first the vertex b'' of $A_j(ab_j)$, where $1 \leqq j \leqq$

$k-1-l$, $j \neq i$, or of $A_i(ab')$. We then follow B_j or B_k, as the case may be, further towards a, forming more of the new C'_j or C'_k. Again, we are in difficulty only if we reach a vertex of A_s, $1 \leqq s \leqq k-1-l$, lying between a and the vertex b_s (or b' or b''). In this case, we follow B_s (or B_k or B_i) further towards a. The process must eventually come to an end, as there are but a finite number of vertices in the graph. We now have the required chains $C'_1, \cdots, C'_k$.

Putting $k=n-1$, we have the required $n-1$ chains $C'_1, \cdots, C'_{n-1}$, and the proof for case 1 is complete.

Case 2. d is on one of the chains $A_1, \cdots, A_n$, say A_n. Let A_n now stand for only that part of the former A_n lying between a and d. Exactly as we before found $n-1$ distinct chains from d to a which did not touch one of the chains $A_1, \cdots, A_n$ (which chain we called A_n), we now find a similar set of chains. If, first, A_n is a chain we have not touched, it forms the required n-th chain from d to a. If A_i, $i \neq n$, was a chain we have not touched, this chain plus the arc cd forms the required n-th chain. The proof for case 2 is now complete also.

3. *Applications to dual graphs.* In this section we remove the restriction on the graphs we consider, that they contain no 1- or 2-circuits, for during the proofs of the following theorems we may run into graphs with such circuits, and we do not wish to have to avoid them.

A planar graph cannot be more than quintuply connected. For suppose there were a sextuply connected graph. Then there is such a graph containing no 2-circuits. By theorem 5, each vertex is on at least six arcs. But this cannot be, as is easily seen from Euler's formula.* However, there are quintuply connected planar graphs, for instance, the dual graph of the dodecahedron (containing twelve vertices, each of which is on five arcs).

Dual graphs may not have the same connectivity. (The dual of the graph just mentioned is but triply connected, each vertex being on but three arcs.) However, we might define a certain notion, the connectivity of a graph *in the large.* Then to a certain extent, we have the theorem that *dual graphs have the same connectivity in the large.* This will be understood on referring to theorem 9 below. We will not go further into the matter here.

THEOREM 8. *Let G and G' be connected† dual graphs. Let a be a*

* This is the equivalent of the theorem that a map on a sphere contains at least one region with less than six sides.

† The hypothesis that the graphs be connected is obviously unnecessary. We include it merely for convenience.

vertex of G, and let A be the arcs on a. Let A', the corresponding arcs of G', form a circuit. Let b' and c' be two distinct vertices of this circuit, dividing it into the two chains B' and C'. Let B and C be the corresponding arcs of G. Form G'_1 from G' by letting b' and c' coalesce into the vertex a'. Form G_1 from G by replacing a by the two vertices b and c, and letting the arcs B end on b, and the arcs C, on c. Then G_1 and G'_1 are duals, preserving the correspondence between their arcs.

Let R, R', r, r', etc., stand for the ranks of G, G', H, H', etc., respectively. Letting b' and c' coalesce in G' reduced the number of vertices by one, and hence

$$R'_1 = R' - 1.$$

Let H_1 be any subgraph of G_1, and let H'_1 be the complement of the corresponding subgraph of G'_1. We must show that

$$r'_1 = R'_1 - n_1.$$

Let H and H' be the subgraphs of G and G' containing the same arcs as H_1 and H'_1, respectively. Then, as G and G' are duals,

$$r' = R' - n.$$

Case 1. b' and c' are in different connected pieces in H'. Then b and c are in the same connected piece in H_1. For add to H' every arc of G' we can without connecting b' and c', forming the graph I'. Let D' be the remaining arcs of G', and let D be the corresponding arcs in G. Then D is the complement in G of the arcs corresponding to I' in G'. I' is in two connected pieces, while adding any arc of D' renders the graph connected. Therefore, as G' is a dual of G, the arcs D are of nullity 1, while dropping out any one of them reduces the nullity to 0. Hence, by N, theorem 9, D is a circuit. As b' and c' are not connected in I', at least one of the arcs of B' and at least one of the arcs of C' are missing from I', and therefore at least one of the arcs of B and at least one of the arcs of C are present in D. In forming G_1 from G, the circuit D is broken at the vertex a, and becomes a chain D_1 in G_1. As D_1 contains still an arc of B and an arc of C, D_1 joins b and c. As I' contains every arc of H', D_1 is contained in H_1. Hence b and c are connected in H_1.

In forming H'_1 from H' by letting b' and c' coalesce, the numbers of vertices and of connected pieces are each reduced by 1, and hence

$$r'_1 = r'.$$

In forming H from H_1 by letting b and c coalesce, the the number of vertices

is reduced by 1, while the number of connected pieces and of arcs remains the same. Hence

$$n_1 = n - 1.$$

These equations with the equations above give

$$r'_1 = R'_1 - n_1,$$

as required.

Case 2. b' and c' are connected in H'. Then b and c are not connected in H_1. For suppose there were a chain D_1 joining them. The arcs of this chain form a circuit D in G, containing one arc of B and one arc of C. Let D' be the corresponding arcs of G', and let I' be the complement of D'. Then as G' is a dual of G, I' is in two connected pieces, while adding any arc of D' connects the graph. Let $d'e'$ be the arc of B' contained in D', where d' is nearest b', e' is nearest c'. (d' may be b', and e' may be c'). As D' contains no other arc of B', b' and d' are connected in I', as are c' and e'. Adding the arc $d'e'$ to I' connects I'. Therefore d' and e', and therefore also b' and c', were formerly not connected. As I' contains every arc of H', b' and c' are not connected in H' either, a contradiction.

As b' and c' are connected in H', letting them coalesce reduces the rank of H', that is,

$$r'_1 = r' - 1.$$

As b and c are not connected in H', letting them coalesce does not alter the rank or nullity of H_1, that is,

$$n_1 = n.$$

These equations with the first equations give

$$r'_1 = R'_1 - n_1$$

again, as required.

THEOREM 9. *Let G be a non-separable planar graph with the following properties. If the vertices $a_1, a_2, \cdots, a_n$ and their arcs are dropped out, the resulting graph is in two connected pieces, H_1 and H_2, while no proper subset of these vertices, if dropped out, disconnect the graph. Either H_1 is of nullity > 0 or one of the vertices a_i is joined to H_1 by at least two arcs, and the same is true of H_2. Then if G' is a dual of G, G' can be disconnected by dropping out a corresponding set of n vertices and their arcs.*

A definition of a "corresponding set of vertices" is given during the proof.

Let A_i be the arcs on a_i, let $A_i(H_1)$ be those arcs joining a_i to vertices in H_1, let $A_i(H_2)$ be those joining a_i to vertices in H_2, and let $A_i(A)$ be

those joining a_i to other vertices of the set $a_1, \cdots, a_n$, $i = 1, \cdots, n$. There are arcs in $A_i(H_1)$ and $A_i(H_2)$ for each i. For suppose for instance there were no arcs in $A_1(H_1)$. Then if we drop out the vertices $a_2, \cdots, a_n$ and their arcs from G, a_1 is now connected only to H_2, and hence H_1 and H_2 are not connected, contrary to hypothesis. Let A'_i, $A'_i(H_1)$, $A'_i(H_2)$, $A'_i(A)$ be the arcs of G' corresponding to A_i, $A_i(H_1)$, $A_i(H_2)$, $A_i(A)$ respectively, $i = 1, \cdots, n$.

The arcs A_i, if dropped out, disconnect G.† Hence the arcs A'_i form a circuit, $i = 1, \cdots, n$. Let us form G^*_1 from G by dropping out the arcs $P = A_1(H_1), A_2(H_1), \cdots, A_n(H_1)$. G^*_1 is in two connected pieces, one of them being H_1, the other being H_2 together with the vertices $a_1, \cdots, a_n$. Putting back any of these arcs connects the two pieces. Hence the corresponding arcs of G' form a circuit P'.

The arcs $A'_i(H_1)$ for any i form a chain. For suppose not. Then for some i, $A'_i(H_1)$ is composed of k chains $C'_1, \cdots, C'_k$. The rest of the circuit A'_i is composed of k chains $D'_1, \cdots, D'_k$, and each chain D'_j joins some two chains of the set $C'_1, \cdots, C'_k$. We formed G^*_1 from G by dropping out the arcs P. Let us form G^*_2 by dropping out also the arcs corresponding to the arcs of D'_1. As D'_1 has its end vertices in P', P' and D'_1 are together of nullity 2. Hence G^*_2 must be in three connected pieces. But this is not so. For the arcs of $A_i(H_2)$ and $A_i(A)$ all join a_i to that part of G^*_1 containing H_2, and D_1 contains but a part of these arcs. Hence a_i is still joined to the rest of this part of G^*_1, and G^*_2 contains the same number, 2, of connected pieces as G^*_1, a contradiction. This proves the statement.

Let us interchange the names of the chains $A'_1(H_1), A'_2(H_1), \cdots, A'_n(H_1)$ so that they occur in that order in P', and rename the vertices $a_1, a_2, \cdots, a_n$ and their arcs accordingly. Let a'_1 be the vertex joining $A'_1(H_1)$ and $A'_2(H_1), \cdots$, let a'_{n-1} be the vertex joining $A'_{n-1}(H_1)$ and $A'_n(H_1)$, and let a'_n be the vertex joining $A'_n(H_1)$ and $A'_1(H_1)$. These vertices form a *corresponding set of vertices* to the vertices $a_1, \cdots, a_n$. If we map the graphs together on a sphere ‡ as described in N, theorem 30, we can draw a closed curve which passes successively through the vertices $a_1, a'_1, a_1, a'_2, \cdots, a_n, a'_n$, and touches no other vertices or arcs. We shall show that $a'_1, \cdots, a'_n$, if dropped out, disconnect G'.

Form G'_{n-1} from G' by letting a'_n and a'_{n-1} coalesce. Form G_{n-1} from G

† The resulting graph is in just two pieces as a_i is not a cut vertex of G, G being non-separable.

‡ So that each arc of one graph cuts the corresponding arc of the other, and otherwise the graphs do not intersect.

by replacing the vertex a_n by the two vertices $a_n(H_1)$ and $a_n(H_2)$, letting the arcs $A_n(H_1)$ end on $a_n(H_1)$, and letting the arcs $A_n(H_2)$ and $A_n(A)$ end on $a_n(H_2)$. By theorem 8, G_{n-1} and G'_{n-1} are duals. Form G'_{n-2} from G'_{n-1} by letting a'_{n-1} and a'_{n-2} coalesce. Form G_{n-2} from G_{n-1} by replacing the vertex a_{n-1} by the two vertices $a_{n-1}(H_1)$ and $a_{n-1}(H_2)$, and replacing the arcs on a_{n-1} as above.* Then G_{n-2} and G'_{n-2} are duals. Continuing, we have finally the dual graphs G_1 and G'_1, where $a'_1, a'_2, \cdots, a'_n$ have all coalesced to form a single vertex in G'_1, and a_i has been replaced by the two vertices $a_i(H_1)$, $a_i(H_2)$, $i = 2, 3, \cdots, n$, in G_1.

Let I_1 be H_1 together with the vertices a_1, $a_2(H_1), \cdots, a_n(H_1)$, and the arcs $A_1(H_1), \cdots, A_n(H_1)$. Define I_2 similarly. Then I_1 and I_2 have but the vertex a_1 in common, and together they form G_1. Let $I_{11}, \cdots, I_{1k}$ be the components of I_1, (there may be but one), and let $I_{21}, \cdots, I_{2l}$ be those of I_2. I_1 is of nullity > 0. For if H_1 is of nullity > 0, I_1 is, as it contains H_1. Otherwise, by hypothesis, there are at least two arcs in one of the sets of arcs $A_i(H_1)$. Let these be the arcs a_ib and a_ic in G. b and c are joined by a chain in H_1, as H_1 is connected; this chain together with the arcs a_ib and a_ic forms a circuit in I_1. The statement now follows from N, theorem 4. Similarly, I_2 is of nullity > 0. Thus one of the components I_{1i} of I_1, and one of the components I_{2j} of I_2, is of nullity > 0 by N, theorem 13.†

Let $I'_{11}, \cdots, I'_{2l}$ be the subgraphs of G'_1 whose arcs correspond to the arcs of $I_{11}, \cdots, I_{2l}$. By N, theorem 25,‡ $I'_{11}, \cdots, I'_{2l}$ are the components of G'_1, and I'_{11} is a dual of $I_{11}, \cdots, I'_{2l}$ is a dual of I_{2l}. As I_{1i} and I_{2j} are of nullity > 0, I'_{1i} and I'_{2j} are of rank > 0, and they contain therefore no 1-circuits. Drop out any 1-circuits there may be in G'_1. The resulting graph J' is separable, as it contains both I'_{1i} and I'_{2j}. As it is connected, it has a cut vertex x'. J' is formed therefore of two graphs J'_1 and J'_2, each containing an arc and hence at least two vertices, which have only the vertex x' in common. If x' is dropped out of J', J' is disconnected. Hence if x' is dropped out of G'_1, forming the graph K'_1, K'_1 is not connected.

x' is the vertex a'_1. For suppose it were not. Form K' from G' by dropping out x' and its arcs. As G is doubly connected, we see easily that G' is also, by N, theorem 26,§ and hence K' is connected. K'_1 is formed from K' by letting the vertices $a'_1, \cdots, a'_n$ coalesce, and hence K'_1 also is con-

* It is easily seen that a_{n-1} is not a cut vertex of G_{n-1}, and hence A'_{n-1} is still a circuit in G'_{n-1}.

† P, theorem 3.

‡ P, theorem 13.

§ P, theorem 14.

nected, a contradiction. Now drop out the vertices $a'_1, \cdots, a'_n$ and their arcs from G'. We are left with the graph K'_1 which is not connected, and the theorem is proved.

THEOREM 10. *A dual G' of a triply connected graph G containing no 1- or 2-circuits is a similar graph.*

G' contains no 1- or 2-circuits; for if it did, dropping out the corresponding one or two arcs of G would disconnect this graph, contrary to theorem 5.

G' contains at least four vertices. For if it contained less than four vertices, it would contain at most three arcs, and would be of nullity $\leqq 1$, and G would be of rank $\leqq 1$, and would contain but two vertices.

Finally, G' cannot be disconnected by dropping out but two vertices. For suppose it could. If dropping out a'_1 and a'_2 and their arcs leaves the two parts H'_1 and H'_2, then, as G cannot be similarly disconnected, either a'_1 and a'_2 are each joined to H'_1 by but a single arc, or they are each joined to H'_2 by but a single arc, say the first, by theorem 9. The two arcs joining a'_1 and a'_2 to H'_1, if dropped out, disconnect G'. Hence the corresponding two arcs of G form a circuit. But this is contrary to the hypothesis that G contains no 2-circuit. Thus G' is triply connected.

THEOREM 11. *A triply connected planar graph containing no 1- or 2-circuit has a unique dual.*

For suppose G' and G'' were both duals of the triply connected graph G. Let H' be any subgraph of G' of nullity 1. Let H'' be the corresponding subgraph of G'' (the correspondence being given through the graph G), and let H be the complement of the corresponding subgraph of G. Then

$$r = R - n',$$
$$r = R - n'',$$

and hence

$$n'' = n',$$

and H'' is of nullity 1. Similarly, if H'' is any subgraph of G'' of nullity 1, the corresponding subgraph H' of G' is of nullity 1. Moreover, by theorem 10, G' and G'' are both triply connected, and neither contains any 1- or 2-circuits. Therefore, by theorem 3, G' and G'' are congruent; that is, the dual of G is unique.

HARVARD UNIVERSITY.

ON THE DECOMPOSABILITY OF CLOSED SETS INTO A COUNTABLE NUMBER OF SIMPLE SETS OF VARIOUS TYPES.

By G. T. Whyburn.

1. In this paper we shall consider the following general problem: *Given any class K of closed point sets, under what conditions is a compact metric space C expressible as the sum of a countable number of K-sets?*

It is obvious that this problem is a much more general formulation of the well known problem * in the theory of curves of finding necessary and sufficient conditions in order that a given curve should be the sum of a countable number of arcs. For, in order to obtain the solution to this special problem from the solution to our general problem we have only to take K to be the class of simple arcs and take C to be the curve. This particular problem was considered by Menger in 1926, who formulated it for *regular curves* and gave the solution for the special case of acyclic curves. The same problem in slightly more general form was considered by the present author in 1929 † and was reduced, for the case of any locally connected continuum, to the same problem concerning the cyclic elements of that continuum. This reduction gave at once the solution to the problem in a number of special cases including those solved by Menger. These remarks suffice to show the point of view from which this problem has been attacked, i. e., the already known classes of continua have been examined in turn to determine if they do or do not have the property of being expressible as the sum of a countable number of arcs or to find conditions under which they do have this property. Since it was found that certain curves of a particular type do have this property whereas other curves of this same type and of apparently no greater complexity do not have this property, the possibility of a general solution to the problem seemed remote. However if, as we shall do in this paper, we concede at once that the continua which have the property of being the sum of a countable number of arcs (or, in general, of K-sets) form a distinct class by themselves which is not to be identified with any of the known types

* See, for example, K. Menger, *Mathematische Annalen*, Vol. 96 (1926), pp. 575-582, and *Jahresbericht der Deutschen Mathematischer-Vereinigung*, Vol. 35 (1926), p. 146.

† See *Fundamenta Mathematicae*, Vol. 14 (1929), pp. 103-106.

in our present classification of curves, and question ourselves as to what type of complexity one can get in a continuum which is formed by adding together two, three, n, or a countably infinite number of arcs (or, in general, of K-sets) we are led at once to those properties which are characteristic for the decomposability of any closed set into a countable number of arcs or of K-sets. Thus we find the solution not only to the arc-problem but to the general problem for any class K of closed sets.

There are many other interesting cases (i. e., choices of the class K) to our general problem and its solution to be given below in §§ 3 and 4 besides that in which K is the class of simple arcs. Some of these will be discussed below in § 5. Notably, for a reason to be mentioned later, the case where K is the class of acyclic curves is of especial interest and importance. Also closed sets which are the sum of *finite numbers* $2, 3, \cdots n, \cdots$ of K-sets have interesting properties which will be discussed in § 6.

Before proceeding with the solution to our problem we introduce a new notion defining sets which may be regarded as

2. *Generalized derived aggregates.* Let C be any compact metric space, let A be any closed subset of C and let K be any class of closed sets. Then by the K-derivative $K(A)$ of A is meant the set of all points x of A such that for no neighborhood U of x is $A \cdot U$ contained in any K-set in C. Set

$$\begin{gathered} A_K^1 = K(A), \\ A_K^2 = K(A_K^1), \\ \cdots\cdots\cdots \\ A_K^n = K(A_K^{n-1}), \\ \cdots\cdots\cdots \end{gathered}$$

in general, $A_K^\alpha = K(A_K^{\alpha-1})$ or $= \prod_{\beta<\alpha} K(A_K^\beta)$, according as the ordinal number α does or does not have an immediate predecessor.

For example, if we take K to be the class of all single points, then for any closed set A, the K-derivatives of A are equal respectively to the ordinary successive derived aggregates of the set A, i. e., $A_K{}^\alpha = A^{(\alpha)}$ for every α. Other examples are as follows. Let H, K, and L denote respectively the classes of simple arcs, simple closed curves, and acyclic curves. Then if C is a circle plus two diameters xoy and poq, $C_H{}^1 = x + y + o + p + q$ while $C_K{}^1 = C_L{}^1 = C_H{}^2 = 0$. If C is the curve obtained by adding to a rectangle $abcd$ a sequence of its altitudes, a_1b_1, a_2b_2, a_3b_3, $\cdots$ converging to ab, then $C_H{}^1 = ab + (a_1 + b_1) + (a_2 + b_2) + \cdots$, $C_H{}^2 = 0$, $C_K{}^1 = C_L{}^1 = ab$, and $C_K{}^2 = C_L{}^2 = 0$. If we add on to this curve a sequence of altitudes con-

verging to each of the altitudes a_ib_i, clearly we get a curve whose third but not second H-, K-, and L-derivatives are vacuous. These examples perhaps suffice to indicate how the K-derivatives may be used to describe or characterize various stages of complexity in the structure of a continuum.

It follows at once from the definition that all the K-derivatives of a closed set A are themselves closed sets and that they decrease monotonically with successive derivations, i. e., $A_K{}^{\alpha} \supset A_K{}^{\alpha+1}$.

3. THEOREM. *In order that the compact metric space C should be the sum of a countable number of K-sets, where K is any given class of closed point sets, it is necessary and sufficient that $C_K{}^{\beta} = 0$ for some ordinal number β of the first or second class.*

To prove the sufficiency of the condition, let us suppose that $C_K{}^{\beta} = 0$, where β is an ordinal number of the first or second class. For symmetry let us set $C = C_K{}^0$. Now each point x of $C - C_K{}^1$ is contained in a neighborhood Q_x such that some K-set K_x in C contains $C \cdot Q_x$. By the Lindelöf Theorem, a countable collection $[Q_i{}^0]$ of the neighborhoods $[Q_x]$ covers $C - C_K{}^1$, and thus $C - C_K{}^1$ is contained in the sum G_0 of the elements of the corresponding countable collection of K-sets $[K_i{}^0]$. Likewise $C_K{}^1 - C_K{}^2$ is contained in the sum G_1 of a countable number of K-sets, $K_1{}^1, K_2{}^1, K_3{}^1, \cdots$ in C. In general, for each $\alpha < \beta$, $C_K{}^{\alpha} - C_K{}^{\alpha+1}$ is contained in the sum G of a countable number of K-sets $K_1{}^{\alpha}, K_2{}^{\alpha}, \cdots$ in C. But it is seen at once that

$$C = \sum_{0 \leq \alpha \leq \beta} (C_K{}^{\alpha} - C_K{}^{\alpha+1}) \subset \sum_{0 \leq \alpha \leq \beta} G_{\alpha} = \sum_{0 \leq \alpha \leq \beta} \sum_{n=1}^{\infty} K_n{}^{\alpha}$$

and, there being only a countable number of ordinals $\alpha < \beta$, we thus have expressed C as the sum of a countable number of K-sets.

The condition is also necessary. For suppose $C = K_1 + K_2 + K_3 + \cdots$, where each K_i is a set of class K, and suppose, contrary to our theorem, that $C_K{}^{\alpha}$ is non-vacuous for all α's of the first and second classes. Then since there are uncountably many of the sets $C_K{}^{\alpha}$, it follows by the well known theorem that (in the space we are considering) any well-ordered monotone decreasing sequence of distinct closed sets is countable that for some ordinal β of the first or second class, $C_K{}^{\beta} = C_K{}^{\beta+1}$. In other words, for some β, $K(C_K{}^{\beta}) = C_K{}^{\beta}$ and hence $C_K{}^{\beta}$ is its own K-derivative. Now clearly there exists a point x_1 of $C_K{}^{\beta}$ which does not belong to K_1; and since K_1 is closed, a neighborhood R_1 of x_1 exists such that $\bar{R}_1 \cdot K_1 = 0$. Likewise, since, by virtue of the definition of K-derivative, $R_1 \cdot C_K{}^{\beta}$ cannot be contained in any K-set in C, there exists a point x_2 in $R_1 \cdot C_K{}^{\beta}$ not belonging

to K_2. There exists a neighborhood R_2 of x_2 such that $\bar{R}_2 \subset R_1$ and $\bar{R}_2 \cdot K_2 = 0$. Similarly in $R_2 \cdot C_K{}^{\beta}$ there exists a point x_3 not belonging to K_3 and there exists a neighborhood R_3 of x_3 such that $\bar{R}_3 \subset R_2$ and $\bar{R}_3 \cdot K_3 = 0$, and so on. Continuing in this way we obtain a sequence of neighborhoods $R_1, R_2, R_3, \cdots$ in C such that for each n, $\bar{R}_n \subset R_{n-1}$ and $\bar{R}_n \cdot \sum_1^n K_i = 0$. But then $\prod_1^\infty \bar{R}_i \cdot \sum_1^\infty K_i = \prod_1^\infty \bar{R}_i \cdot C = 0$, contrary to the fact that $\prod_1^\infty \bar{R}_i$ is not vacuous and is a subset of C. This contradiction proves our theorem.

4. *An alternate form.* We can state our proposition just proved in the following equivalent form.

In order that a compact metric space C should be the sum of a countable number of K-sets it is necessary and sufficient that each subset N of C contain at least one point x for some neighborhood R of which it is true that $N \cdot R$ is contained in some K-set in C.

For if this condition is satisfied, then by choosing N to be the successive K-derivatives of C we see that these sets $C_K{}^{\alpha}$ must be distinct sets and since they are well-ordered and monotone decreasing, they must be countable in number. Thus they must be vacuous after a certain β of the first or second class, and by our theorem in § 3, C is the sum of a countable number of K-sets. If, on the other hand, C is the sum of a countable number of K-sets, then $C_K{}^{\beta} = 0$ for some β of the first or second class; but it is easily seen that if any subset N of C failed to contain a point x such that the part of N is some neighborhood of x is contained in a K-set in C, then N would belong to every set $C_K{}^{\beta}$ for all ordinal numbers β.

5. *Applications. Some particular choices for K.*

(a). In general, for any given class K, let us call a continuum ($=$ a self-compact, connected, metric space) C which is the sum of a countable number of K-sets a *K-sum*. Then from § 4 it follows at once that *in any K-sum C there exists an everywhere dense set of open sets each of which is contained in some K-set in C.*

(b). If we take K to be the class of simple arcs, then § 4 tells us that in order that a continuum C should be an arc-sum it is necessary and sufficient that every subset N of C contain a point x for some neighborhood R of which it is true that $N \cdot R$ is a subset of some arc in C. Thus we have a solution to the arc-sum problem. In particular, if we take $N = C$, we see that any arc-sum C must contain open sets which are contained in arcs in C

and thus it must contain "free-arcs," i. e., arcs ab such that no point of $ab - (a+b)$ is a limit point of $C - ab$. Indeed, by (a) it follows that *in any arc-sum the free arcs must be everywhere dense.* Suppose we call a continuum in which the free arcs are everywhere dense a *free arc continuum.* Then we have that *any arc-sum is a free arc continuum.*

We note here, in answer to a problem proposed by Menger (*loc. cit.*), the fact that *any arc-sum A is the sum of the elements in a null sequence of arcs,* i. e., $A = \sum_1^\infty B_n$, where each B_n is an arc and $\delta(B_n) \to 0$ as $n \to \infty$. Clearly this is the case; for if $A = \sum_1^\infty A_n$, where each A_n is an arc, then for each n, A_n is the sum of a finite number of arcs A_n^i each of diameter $< 1/n$; and obviously we can arrange these arcs $[A_n^i]$ into a null sequence. Indeed, similar reasoning yields at once the result that *if K is any class of closed sets such that each K-set is itself the sum, for each $\epsilon > 0$, of a finite number or of a null sequence of K-sets of diameter $< \epsilon$, then every K-sum is the sum of the elements in a null sequence of K-sets.*

(c). Suppose now we take K to be the class of acyclic curves, i. e., the locally connected continua containing no simple closed curves. Then of course §§ 3 and 4 give us necessary and sufficient conditions for a continuum C to be an acyclic curve-sum. Now in case C itself is a locally connected continuum, this choice of K is of special significance, because of the fact * that any closed subset of such a set C which is homeomorphic with a subset of some acyclic curve *is really contained in an acyclic curve in C.* Thus in this case we can characterize acyclic curve-sums completely in terms of topological properties of their subsets as follows: *A locally connected continuum C is an acyclic curve-sum if and only if every closed subset N of C contains a point x such that some closed neighborood of x in N is homeomorphic with a subset of some acyclic curve* (i. e., has only point- and acyclic curve-components and only a null sequence of components which are not points). Also in this case it follows at once that *any locally connected subcontinuum of an acyclic curve-sum is itself an acyclic curve-sum.* It has been remarked by Menger (*loc. cit.*) that the corresponding proposition about arc-sums is not valid. Thus it appears that in many respects a detailed study of acyclic-curve-sums would be more fruitful than one of arc-sums.

(d). If we choose K to be the class of free arc continua, we get at once

* This is established for C in the plane by H. M. Gehman [See *Transactions of the American Mathemathical Society,* Vol. 29 (1927), p. 560], and has been extended to general metric locally connected continua by L. Zippin.

from § 4 or from (a) that the free arcs are dense in any free arc continuum-sum and hence *any free arc continuum-sum is itself a free arc continuum.*

(e). As a final application, let us choose K to be the class of continua without continua of condensation, which we shall call w-continua. Then since any w-continuum is * a *free arc continuum,* it follows as a corollary to (d) that any w-continuum-sum is a free arc continuum. Furthermore it is readily proved with the aid of some elementary results of dimension theory that *any subset of a w-continuum-sum which contains no arc is of dimension zero.* Thus any 1-dimensional subset (e. g., any connected subset) of a w-continuum-sum must contain an arc. In particular (since an arc is a w-continuum), any arc-sum has the property just mentioned.

6. *The finite cases.* In conclusion we make some remarks concerning the cases in which $C_K{}^n = 0$ for finite integers n. It might be supposed, by analogy with our result in § 3, that the vanishing of $C_K{}^n$ for a finite integer n comprises a necessary and sufficient condition that C be the sum of n K-sets. However, we shall see at once that in general this is not the case. That the condition is not sufficient is seen for example, in the continuum C obtained by adding on to the unit interval a perpendicular ordinate of length $1/q$ for each point p/q where p and q are integers prime to each other and $p < q$; for in this case, clearly C is not the sum of any finite number of arcs, whereas its second arc-derivative vanishes (i. e., $C_K{}^2 = 0$, where K is the class of simple arcs). It will be shown below, however, that this condition is necessary.

In general, *if C is the sum of a finite number of K-sets, then for any integer n, every point $C_K{}^n$ belongs to at least $n + 1$ of these sets.* For clearly every point of $C_K{}^1$ belongs to at least 2 of the K-sets; and assuming that for an integer $k - 1$, every point of $C_K{}^{k-1}$ belongs to at least k of the sets, it follows that every point of $C_K{}^k$ belongs to at least $k + 1$. For if not, then some point x of $C_K{}^k$ belongs to at most k of the sets, say to $K_1, K_2, \cdots, K_k$; hence some neighborhood U of x in C is a subset of $K_1 + K_2 + \cdots + K_k$ and thus $C_K{}^{k-1} \cdot U \subset K_1 + K_2 + \cdots + K_k$; but then since by supposition every point of $C_K{}^{k-1}$ belongs to at least k of the sets, it follows that $C_K{}^{k-1} \cdot U \subset K_1 \cdot K_2 \cdots K_k$; thus the part of $C_K{}^{k-1}$ in a neighborhood of x is a subset of K-set in C, (e. g., of K_1), contrary to the fact that $x \subset C_K{}^k$. Therefore, by induction, it follows that for any n, every point of $C_K{}^n$ belongs to at least $n + 1$ of the sets.

Now, in particular, if C is the sum of n K-sets, every point of $C_K{}^{n-1}$

* See Urysohn, *Verhandelingen der Akademie te Amsterdam,* Vol. 13 (1927), No. 4.

is contained in all n of these sets and hence C_K^{n-1} is contained in their product. It follows from this that $C_K^n = 0$, because C_K^{n-1} is a subset of some single K-set in C. This also follows because if C_K^n contained a point x, then x would have to belong to $n + 1$ of the sets, whereas there are only n sets in all. Therefore we have proved that *In order that a set C be the sum of n K-sets, it is necessary that C_K^{n-1} be a subset of a K-set in C and that $C_K^n = 0$.*

Although, as we have seen, this condition is not sufficient for classes K in general, nevertheless it seems possible that it may be sufficient for certain particular classes K. Notably when K is the class of acyclic curves does this seem possibly to be the case. It would be interesting to study the finite cases further and find necessary and sufficient conditions that a set C be the sum of n K-sets in general and for particular choices of K.

THE JOHNS HOPKINS UNIVERSITY.

GENERALIZATION OF A THEOREM DUE TO C. M. CLEVELAND.†

By Leo Zippin.‡

Recently, C. M. Cleveland has announced the theorem that: if, in the plane, M is a bounded continuous curve which contains no domain and K is a totally disconnected closed subset of M such that $M - K$ is connected, there exists an acyclic continuous curve T containing K and such that 1) all endpoints of T belong to K, 2) the point set $M \cdot T$ is disconnected, 3) $M - M \cdot T$ is connected.§ It is, of course, clearly understood that T belongs to the plane of M. It was the suggestion of Professor R. L. Moore to extend this result to euclidean spaces of higher dimension. It is to be expected that the theorem will admit of considerable elaboration in higher spaces, and we give two distinct versions both considerably stronger than the original form; however, the methods of this paper do not admit of application in the plane case, and the results are not generally true in the plane. Our theorems generalize, also, certain results of R. L. Moore relating to "paths that do not separate a given continuous curve," and results due to Zarankiewicz. The arguments of this paper rest principally on well known theorems of Dimension-Theory, and the local arcwise connectedness of generalized continuous curves C, i. e., complete separable metric spaces, connected and locally connected.¶ We give, in particular, a necessary and sufficient condition that a self compact totally disconnected subset of C be the set of endpoints of an acyclic continuous subcurve of C.

Preliminary: Let B be a closed, non-dense subset of a generalized continuous curve C (in the sense above) which does not locally separate C.‖

† Presented to the American Mathematical Society, February 11 (1930). Read at the Easter Meeting, Berkeley, California.

‡ National Research Fellow.

§ See, *Bulletin of the American Mathematical Society*, Abstract, Vol. 36, p. 331. The essential condition that $M - K$ is connected is, by oversight certainly, not there stated.

¶ See R. L. Moore, *Bulletin of the American Mathematical Society*, Abstract, Vol. 33 (1927), p. 141; Synopsis of Boulder Colloquium Lectures (August 1929); the Colloquium Lecture to appear in the regular series. Also, K. Menger, *Monatshefte für Mathematik und Physik*, Vol. 36 (1929), p. 212; N. Aronszajn, *Fundamenta Mathematicae*, Vol. 15 (1930), p. 228; in the same volume, C. Kuratowski, p. 301.

‖ We shall say that a subset V of C locally separates C provided there exists an open connected subset U of C such that $U \cdot V = V'' \subset \bar{V}'' \subset U$, and $U - V''$ is not connected.

There exists a countable set P of points of $C-B$, dense in $C-B$ and therefore dense in C. For every pair p and q of points of P there is an arc pq in $C-B$: for $C-B$ is by hypothesis connected, and is connected im kleinen since B is closed; and being open in C, it is by a theorem due to Alexandroff † in itself a complete metric space.

We suppose well ordered the denumerable set of pairs p and q of points of P. For each pair (p,q) let pq be an arbitrary arc of $C-B$ of diameter less than $X_{pq}+1/n$, where n is the integer associated with the pair (p,q) in the well ordering, and X_{pq} is the lower limit of the set of diameters of the totality of arcs joining p and q in $C-B$. Let K represent the point set which is the sum of all of these arcs.

It is clear that $K\cdot B=0$, that K is connected, and since B is closed, that K is connected im kleinen: for the last, if k is a point of K, there is a d such that no point of C at a distance less than or equal to d from k belongs to B; in symbols, $\bar{S}(k,d)\cdot B=0$. Let n be an integer such that $1/n<d/2$. There is an e such that two points of $S(k,e)$ may be joined by an arc of $S(k,1/n)$ and such that $S(k,e)$ contains no pair of points of P of subscript less than or equal to n. Then $K\cdot S(k,e)$ belongs to a connected subset of $S(k,d)$. If H is any subset of $C-K$ we shall show that H does not even locally separate C. Otherwise there is an open connected subset U of C such that $U\cdot H=H''\subset\bar{H}''\subset U$, and $U-H''$ is not connected. Then H'' contains a closed subset H^* which irreducibly separates U between two points of U.‡ It follows from our condition on B that H^* is not a subset of B and contains a point h not of B. There is an e such that $S(h,e)\subset U-B\cdot U$ and a d such that two points of $S(h,d)$ may be arc-joined in $S(h,e/2)$. Let n be an integer such that $1/n<e/2$. There exist two sequences of points of P, (p_i) and (q_i) belonging to distinct components of $U-H^*$, and converging to h. There exists a pair of points p_j and q_k such that $p_j+q_k\subset S(h,d)$ and the point pair $|(p_j,q_k)|$ is of subscript greater than n in the well ordering of pairs of points of P. The arc p_jq_k of K belongs to $U-H^*$.

With the weaker condition on B that B does not separate C (and other-

† "Sur les ensembles de la première classe et les ensembles abstraits," *Comptes Rendus*, Vol. 178 (1924), pp. 185-187. (That the metric which exhibits the completeness of $C-B$ is in general different from the metric originally associated with C need not detain us, who are concerned merely with the fact of its arcwise connectedness).

‡ Compare G. T. Whyburn, "Irreducible Cuttings of Continua," *Fundamenta Mathematicae*, Vol. 13 (1929), p. 49, Theorem 8: the thought of this proof generalizes readily.

wise K as defined above is not itself connected), it is immediate from the connectedness and "everywhere-density" of K that no subset of $C-K$ separates C. With the stronger condition on B that B does not even locally separate in the weak sense, i. e. if U is an open connected subset of C and $U \cdot B = B^*$ then $U - B^*$ is connected, it follows as above that no subset of $C-K$ locally separates even in the weak sense. It may be of interest to note that for the case that B is vacuous, K defined above includes all cut points and local cut points of C, and that this fact may be employed to yield simple proofs that these sets are at most one dimensional and are F_σ-subsets of C.†

In view of a recent paper by Ayres which has just been received by the author, it has seemed well to interpolate here a justification of the previous statements, especially since the lemma ‡ there employed is not valid for generalised continuous curves: witness the simple example that C is the following point set (in the plane): $0 \leq x \leq 1$, $y = 0$; $0 \leq y \leq 1$, $x = 1/n$, $(n = 1, 2, 3, \cdots$. Here the point $(0, 0)$ is not a cut-point, while the points $(1/n, 0)$ belong to H_1 in Ayres' notation. Now, if the point y of C is a local cut-point of C there is an e such that y is a cut-point of the composant U_y of $S(y, e)$. If y is a point of an arc pq of K, defined above, we may choose e so that p and q are not in $S(y, e)$. On pq there is a pair of points x and z, in order $pxyzq$, such that $(yx - x) + (yz - z) \subset U_y$, and $x + z$ do not belong to U_y. If y fails to separate some point of $yx - x$ from some point of $yz - z$ it cannot separate any pair of points of these respective sets, and there is a point u of U_y separated by y from both of these point sets: in this case there is an arc uy which has only y in common with $pxyzq$, and the set of such points y (by a theorem of Zarankiewicz, from the separability and metricity of C) is at most a countable set Y. If the point y is not a point of Y it must separate points of $yx - x$ from points of $yz - z$ (in U_y for the case of local cut-points) and belongs to H_n where $1/n \leqq e$.§

Suppose now that y' is not a point of H_n, for a given n, and is not a point of Y, but is a point of the arc pq (with the understanding as before that p and q are not in $S(y', 1/n)$, the other case offering no difficulty, and

† G. T. Whyburn (and references there to Zarankiewicz), "Concerning Points of Continuous Curves Defined by Certain im kleinen Properties," *Mathematische Annalen*, Vol. 102 (1929), p. 318, Theorem 8.

‡ W. L. Ayres, "A New Proof of a Theorem of Zarankiewicz," *Fundamenta Mathematicae*, Vol. 16 (1930), p. 134. Compare, for this section, G. T. Whyburn, "On the Structure of Connected and Connected im kleinen Point sets," *Transactions of the American Mathematical Society*, Vol. 32, pp. 926-928.

§ It will be clear that we are defining H_n.

the endpoints p and q being considered separately from the inner points). There is then an arc $x'y'z'$ of pq, such that $x'y'z'$ belongs to $S(y', 1/n)$ excepting for its endpoints. Since y' is not in H_n there is an arc in $U_{y'}$, the component containing y' of $S(y', 1/n)$ joining the point x'' of $x'y' - y'$ to a point z'' of $z'y' - y'$, the arc having these points only in common with $x'y'z'$, and being contained in $S(y', d)$ for some positive number $d < 1/n$. It is readily seen that there is a neighborhood of y' no point y^* of which separates points of the corresponding are $x^*y^*z^*$ in the component U_{y^*} of $S(y^*, 1/n)$, since for points sufficiently near to y', $x^*y^*z^* \supset x'y'z'$ and $U_{y^*} \supset U^0_{y'}$, the component of y' in $S(y', d)$: in other words, H_n defined for pq is closed there. Then the set of local cutpoints of C, these being contained in arcs of K, is the sum of a countable number of point sets, and each of these is the sum of a countable number of closed sets plus a countable point set Y. Finally we add those points of P, end-points of arcs pq, which are local cut-points. The local cut-points of C form an F_σ-subset of C: and an entirely similar argument holds for the set of cut-points. Finally that these sets are at most one-dimensional, is an immediate consequence of their inclusion in a one-dimensional point set.

We turn to the first generalization of Cleveland's Theorem.

THEOREM 1. *If C is a continuous curve, lying in E_n $(n > 2)$, a euclidean space of n-dimensions, and contains no domain in E_n, and B is a closed and totally disconnected subset of E_n such that $B'' = B \cdot C$ does not separate (locally separate) C, then there exists a ray L in E_n such that 1) $L \supset B$, 2) $L \cdot C = B^0$ is totally disconnected, 3) B^0 does not separate (locally separate) C, 4) $B^0 - B \subset \bar{C}^{n-1}$, where C^{n-1} denotes the points of C which are of dimension at least $n - 1$ in C,† 5) the endpoint of L is an arbitrary point of B, 6) if B is bounded, L is an arc with arbitrary endpoints in B.*

There exists a G_δ-subset G of C such that G is of dimension zero, $C - G$ is an F_σ of dimension at most $n - 2$, and every point of G is a point of $\bar{C}^{n-1}$.‡ From our preliminary considerations there exists in C a one-dimensional F_σ-subset K, such that no subset of $C - K$ separates (locally separates) C. The set $K + (C - G)$ is an F_σ-subset of C of dimension $n - 2$ at most.§ The complement in C of $K + (C - G)$ is a subset G''

† Dimension is to be understood throughout in the Menger-Urysohn sense.

‡ This follows from a slight modification of the argument given by Menger, *Dimensiontheorie*, p. 110, Teubner (1928).

§ *Ibid.*, p. 93.

of G. The complement in E_n of $K + (C - G)$ is a G_δ-subset M''. It is readily seen that $M'' + B$ is a G_δ-subset M of E_n.

That there is in E_n a ray L'' which contains B is not difficult to establish; it is most easily deduced from an interesting theorem due to G. T. Whyburn,† as follows: If B is unbounded, we invert E_n with respect to a center of inversion v not belonging to B. Then $v + B^*$, the image of B, is bounded, closed, totally disconnected. There is an arc x^*v in E_n, where x^* is an arbitrary point of B^*, containing $B^* + v$. The image of x^*v in E_n is a ray L'' whose endpoint is x, the image of x^*. In the case that B is bounded, the theorem of Whyburn gives immediately an arc xy in E_n containing B, and the argument of the sequel may be applied directly to it. Since B is closed and totally disconnected, the complement of B in L'' is a countable set of intervals with endpoints a and b in B. Let ab be such an interval. Let e be an arbitrary positive number less than 1. It is readily seen that there exists an infinite chain of spheres of E_n of order type $w^* + w$, i. e., $\cdots S_{-n}, S_{-(n-1)}, \cdots S_0, S_1, \cdots S_n, \cdots$, such that the diameter of S_j is at most the numerical value of e/j (at most e, if $j = 0$), $\bar{S}_i$ and $\bar{S}_j$ have a common point if and only if i and j are consecutive, $S_i \cdot ab$ is not vacuous for any i, and $S_i \cdot (L'' - ab) = 0$ for every i. Since the complement of M in E_n is of dimension at most $n - 2$, M is everywhere dense in E_n and there exists a chain of points of M of order type $w^* \neq w$; $\cdots x_{-n}, \cdots x_0, \cdots, x_n, \cdots$ such that $x_0 \subset S_0$, and $x_i \subset S_i \cdot S_{i+1}$ if i is positive, and $x_i \subset S_i \cdot S_{i-1}$ if i is negative.

Now let x be any point of M and S any sphere of E_n containing x. Since the complement of M in E_n is of dimension at most $n - 2$, $M \cdot S$ is connected.‡ Likewise, it is locally connected. Being a G_δ-subset of E_n (it will be recalled that M is a G_δ-subset) it is, by a theorem of Alexandroff, in itself a complete space, and therefore arcwise connected. The sum of the infinite set of arcs x_ix_j for every pair of *consecutive points* of $\cdots x_{-n}, \cdots, x_n, \cdots$, where x_ix_j is an arc of $S_i \cdot M$ if i is positive and of S_{i+1} if i is negative, plus the points a and b, is a continuous curve containing a and b, and has a subarc ab of M which has only the points a and b in common with $L'' - ab$ (of L''). It is clear that by ordering the intervals of L complementary to B, and applying the above argument for a sequence of positive numbers e converging to zero to the successively resulting rays we arrive

---

† G. T. Whyburn, "Concerning Continuous Curves Without Local Separating Point," *American Journal of Mathematics*, Vol. 53 (1931), p. 163.

‡ Urysohn, "Sur les multiplicités Cantoriennes," *Fundamenta Mathematicae*, Vol. 8 (1926), p. 310, Lemma. The restriction to closed sets is not essential.

at a ray L contained in M, containing B, and with the endpoint x of L''. Since $L \subset M$ and $M \cdot C \subset G'' + B \subset G + B$, $L \cdot C \subset G + B$. Since B is closed, $G + B$ is zero-dimensional.† Since $L \cdot C = B^0$ is closed and zero-dimensional, it is totally disconnected. Since $B^0 \subset M \cdot C$, $B^0 \subset C - K$ and does not separate (locally separate) C. With this the theorem is proved.

If we suppose, as a very special case, that B consists of two points x and y which do not belong to C, then we have shown that there is an arc xy of E_n whose product with C is totally disconnected and does not even locally separate C. This generalises a result of R. L. Moore.‡ If we suppose that C is any F_σ-subset of E_n of dimension at most $n-2$, and B is any closed and totally disconnected subset of C, the ray L has only B in common with C. This generalizes a theorem due to Zarankiewicz on the accessibility from euclidean 3-space of the point of a one-dimensional continu.§ In particular if a and b are two points of C, and ab is the straight line segment in E_n there exists, by the argument which we gave on the subarc ab of L'' (in the preceding theorem) an arc ab which has only a and b in common with C and whose diameter is the distance from a to b. We shall make an application of this remark to the Moore-Kline Theorem.¶

If M is a Moore-Kline set, bounded, in E_n: i. e. a closed set of points whose maximal connected subsets are points or arcs, and if x in any point of an arc X of M but not an endpoint of X, then x is not a limit point of $M - X$, then M is a one-dimensional F_σ-subset of E_n. For it is the sum of countable set of arcs and a closed totally disconnected point set consisting of endpoints of arcs of M and point-components. There is in E_n an acyclic continuous curve which contains M.‖ There is, by Waszewski-Gehman Theorem an acyclic continuous curve in the plane homeomorphic with the given acyclic continuous curve. By this there is a plane set M'' homeomorphic with M. By the Moore-Kline Theorem there is an arc L'' which contains M''. This imposes an ordering upon the points of M'', and therefore, by the correspondence, upon the points of M. If $a''_i b''_i$ is an arc of L'' whose

---

† K. Menger, *Dimensionstheorie*, p. 114: B is at once a $G\delta$ and F_σ-set.

‡ R. L. Moore, "Concerning Paths That Do Not Separate a Given Continuous Curve," *Proceedings of the National Academy of Sciences*, Vol. 12 (1926), p. 750, Theorem 10.

§ C. Zarankiewicz, "Sur les points de division dans les ensembles connexes," *Fundamenta Mathematicae*, Vol. 9 (1927), p. 166 (3).

¶ R. L. Moore and J. R. Kline, "On the Most General Closed (Plane) Point Set Through Which it is Possible to Pass a Simple Continuous Arc," *Annals of Mathematics*, Vol. 20 (1919).

‖ L. Zippin, "On Continuous Curves and the Jordan Curve Theorem," *American Journal of Mathematics*, Vol. 52 (1930), p. 332, Theorem 2.

endpoints only are points of M'', we shall regard the corresponding point pair a_i and b_i of M as associated. By the homeomorphism and the fact that M and M'' are closed and bounded, it is seen that the distance in E_n between associated pairs of points a_n and b_n approaches zero with increasing n: we suppose the denumerable set of intervals complementary to M'' in L'' arbitrarily ordered. Then M plus the countable set of straight line intervals a_nb_n is a one-dimensional F_σ-subset M^0 of E_n. Replacing the intervals a_nb_n by independent arcs of $E_n \cdot \{M^* - (a_n + b_n)\}$ of the same diameter, as indicated above, we have an arc L which contains M and preserves the order imposed upon M by the homeomorphism with M''. Except for the detail of "order" however, this is a known result.†

To obtain a different version of Cleveland's theorem,‡ and because it may be of interest in itself, we shall turn to a theorem on generalized continuous curves. We shall show, first, that if B is a self-compact totally disconnected subset of a generalized continuous curve (i. e. a complete separable metric space, connected and connected im kleinen) and B does not even locally separate C, then there exists in C an acyclic continuous curve T the set of whose endpoints is B. From the self-compactness of B it readily follows that there exists in C a finite set $U_1, U_2, \cdots, U_n$ of open connected sets of diameter at most 1, such that $\bar{U}_i \cdot \bar{U}_j = 0$, $U_i \cdot B \equiv B_i = \bar{B}_i$ (not vacuous), and $U_i - B_i$ is connected: therefore as we have seen in previous arguments, arcwise connected. If k_i, $1 \leq i \leq n$, is a set of points, $k_i \subset U_i - B_i$, there is in $C - B$ a tree T_1 irreducible about the set k_i. There exists in U_i, $1 \leq i \leq n$, a finite set of open connected subsets U_{ij}, $1 \leq j \leq n_i$, of diameter less than $1/2$, $\bar{U}_{ij} \cdot T_1 = 0$, $\bar{U}_{ij} \cdot \bar{U}_{ik} = 0$, $U_{ij} \cdot B = U_{ij} \cdot B_i \equiv B_{ij} = \bar{B}_{ij}$, and $U_{ij} - B_{ij}$ is arcwise connected. Let k_{ij}, $1 \leq i \leq n$, $1 \leq j \leq n_i$, be a finite set of points, $k_{ij} \subset U_{ij} - B_{ij}$. There is a tree (acyclic continuous curve) T_2 irreducible about the set k_{ij}, and $T_1 \supset T_2 \cdot (C - \Sigma v_i)$; the point set $T_1 \cdot T_2$ is itself a tree: it suffices, for each point k_{ij} (these being arbitrarily ordered) to consider any arc $y_{i1}k_i$ in $U_i - B_i$, and the subarc on this from k_{i1} to the first point on T_1; for succeeding points k_{ij} the subarc from k_{ij} to the first point on T_1 plus the finite set of arcs already added.§ We define inductively a sequence of trees, T_n; the point set compared of points common

† E. W. Miller, *Bulletin of the American Mathematical Society*, Abstract, Vol. 36 (1930), p. 361, 265.

‡ Announced by the author, some time ago, as a theorem for continuous curves.

§ This and other details of the argument parallel closely the proof which has been given of a more restricted theorem in L. Zippin, "A Study of Continuous Curves and Their Relation to the Janiszewski-Mullikin Theorem," *Transactions of the American Mathematical Society*, Vol. 31 (1930), p. 745, Theorem 1.

to all but a finite number of these being itself a locally compact acyclic continuous curve T, $\bar{T} - T = B$, $\bar{T}$ is an acyclic continuous curve (compact) the set of whose endpoints is B.‡

It may be remarked that in order that *every* self-compact, totally disconnected subset B of a generalized continuous curve C be the set of endpoints of an acyclic continuous subcurve of C, it is *necessary* (by the argument above, it is also sufficient) that no self-compact totally disconnected subset H of C even locally separate C. Otherwise there is an open connected subset U of C, $U \cdot H = H'' \subset \bar{H}'' \subset U$, and $U - H''$ is not connected. There is a subset H^0 of H'' which separates U and a point h of H^0 which is a sequential limit point for two distinct sequences (x_i) and (y_i) of $U - H^0$, such that there is no arc $x_i y_j$ in $U - H^0$.§ But then it will be seen that no acyclic continuous curve of C can have the point set $H^0 + (x_i) + (y_i)$ as the set of its endpoints: for some j and k the corresponding arc $x_j y_k$ must have at least one point of H^0 as inner point. Then we have the

THEOREM. *A necessary and sufficient condition that every self-compact totally disconnected subset of a generalized continuous curve C be the set of endpoints of an acyclic continuous curve of C is that no self-compact totally disconnected subset of C even locally separate C.*

We return to a consideration of Cleveland's theorem and prove

THEOREM 2. *If C is a continuous curve in E_n, $n > 2$, and contains no domain, and B is a closed and totally disconnected subset of E_n such that $B'' = B \cdot C$ does not separate (locally separate) C, then there exists an acyclic continuous curve T in E_n such that 1) B is the set of endpoints of T, 2) $T \cdot C = B^0$ is totally disconnected, 3) B^0 does not separate (locally separate) C, 4) $B^0 - B \subset \bar{C}^{n-1}$.*

Consider the point set M defined in the first theorem: it will be recalled that M is a generalized continuous curve containing B whose product with C is a zero-dimensional set of points G'' such that no subset of $G'' + B''$ separates (locally separates) C. The complement of M in E_n is an F_σ-subset K'' (say) of dimension at most $n - 2$. Since $K'' + V$ is at most $(n - 2)$-dimensional, where V is any closed and totally disconnected subset of M, it is readily seen that V does not even locally separate M. If, now, B is

‡ The details are suppressed. The thought of the proof will be found, for example, in H. M. Gehman, "Concerning Acyclic Continuous Curves," *Transactions of the American Mathematical Society*, Vol. 29 (1927), p. 567, Theorem 6; also, reference above.

§ Compare the *preliminary* paragraph.

bounded in E_n it is self-compact and we may apply the immediately preceding theorem to determine an acyclic continuous curve T with the properties of the theorem (these follow as in Theorem 1). If B is unbounded, we invert E_n with respect to a center of inversion v in M. Then B^*, the image of B, plus v is a self-compact set of points: M^*, the image of M, is a generalized continuous curve. This follows from the fact that the image in E_n of the F_σ-set K'' is an F_σ-set K^* because the image of a closed set while not necessarily closed (without the addition of the center of inversion v, which we are reserving to M^*) is always an F_σ. Then there is an acyclic continuous curve T^* in M^* the set of whose endpoints is $B^* + v$, and the image of this in E_n is an unbounded acyclic continuous curve T whose endpoints are the set B. Again, that T has all the properties required of it by the theorem follows as in Theorem 1.

We remark, finally, that if C'' is a continuous curve in E_n and contains domains in E_n, there exists a continuous curve C in C'' which is nowhere dense in E_n and contains every boundary point of C'', and such that $C'' - C$ is a countable set of open spheres (S_n) of $E_n \cdot C''$ the set of whose diameters converges to zero, and $\bar{S}_i \cdot \bar{S}_j = 0$.† This permits us to extend Theorems 1 and 2 to the continuous curves C'' which contain domains in E_n if we replace the condition that the set B^0 defined in those theorems (C'' replacing C) is totally disconnected by this: if X is any arc of points of B^0 and x is an inner point of X, then x belongs to one of the spheres S_n, above. The statements of the resulting theorems will be readily anticipated.

THE UNIVERSITY OF TEXAS,
AUSTIN, TEXAS.

† Compare in J. H. Roberts, "Concerning Non-Dense Plane Continua," *Transactions of the American Mathematical Society*, Vol. 32 (1930), p. 24 (2).

TRITANGENT CIRCLES OF THE RATIONAL BICUBIC.

By Frank Morley and W. K. Morrill.

1. *Introduction.* The general sextic curve in a projective space is the intersection of a quadric surface and a cubic surface. With reference to the quadric surface it is a 3 to 3, or bicubic, correspondence along two generators. We take the quadric as a sphere and isolate a point ∞ of it. It is convenient to speak (inversively) of the sphere (with ∞ marked on it) as a plane. On it are conjugate coördinates x and $\bar{x}$ and the curve is given by a self-conjugate, 3 to 3, relation between x and $\bar{x}$. We call it a bicubic curve,—it is also, projectively, called a tricircular sextic. The tritangent planes of the projective curve are, on the plane, tritangent circles or contact circles.

In connection with a memoir on the tritangent planes by F. P. White, *Proceedings of the London Mathematical Society,* Vol. 30 (1930), p. 347; it occurred to J. H. Grace, *Journal of the London Mathematical Society,* Vol. 5 (1930), p. 121, and to Frank Morley, *American Mathematical Monthly,* Problem No. 3444, Vol. 37 (September 1930), that sets of four tritangent circles must have a common tangent circle. We discuss this, an extension of Hart's theorem,* for the case of the rational curve.

2. *Abel's theorem for sections by a circle.* The equation of a rational bicubic with four double points is given by,

$$x=\frac{\alpha_0+\alpha_1\rho+\alpha_2\rho^2+\alpha_3\rho^3}{\beta_0+\beta_1\rho+\beta_2\rho^2+\beta_3\rho^3}, \tag{1}$$

where x, α_i, β_i are complex quantities and ρ is a real. If we substitute for x and $\bar{x}$ from (1) and its conjugate in the equation $(x-c)(\bar{x}-\bar{c})=r^2$ of a circle we get a sextic $f(\rho)$ in ρ, showing that a circle has with the curve six common points, either actual intersections or common image pairs. If we call a_i, b_i $(i=1,2,3,4)$ the values of the parameter at the four double points, the six parameters of the points on a circle obey the relations,

$$\frac{(a_i-\rho_1)(a_i-\rho_2)(a_i-\rho_3)(a_i-\rho_4)(a_i-\rho_5)(a_i-\rho_6)}{(b_i-\rho_1)(b_i-\rho_2)(b_i-\rho_3)(b_i-\rho_4)(b_i-\rho_5)(b_i-\rho_6)}=k_i,\dagger \tag{2}$$

$$(i=1,2,3,4).$$

* Coolidge, *Circle and Sphere,* p. 43.

† Clebsch, *Vorlesungen über Geometrie der Ebene,* p. 893.

To determine k_1, we suppose the circle to pass through the three double points (a_2, b_2), (a_3, b_3) and (a_4, b_4). Then we have,

$$k_1 = \frac{(a_1-a_2)(a_1-b_2)(a_1-a_3)(a_1-b_3)(a_1-a_4)(a_1-b_4)}{(b_1-a_2)(b_1-b_2)(b_1-a_3)(b_1-b_3)(b_1-a_4)(b_1-b_4)}.$$

Introducing $a_1 - b_1$ in the numerator and denominator and writing,

$$\phi(x) = (x-a_1)(x-b_1)\cdots(x-a_4)(x-b_4),$$

we have $k_1 = -\phi'(a_1)/\phi'(b_1)$, where $\phi'(x)$ represents the derivative of ϕ with respect to x. In like manner we evaluate k_2, k_3, and k_4. We can rewrite the equations (2) in the following way:

$$\frac{(a_1-\rho_1)\cdots(a_1-\rho_6)}{\phi'(a_1)} + \frac{(b_1-\rho_1)\cdots(b_1-\rho_6)}{\phi'(b_1)} = 0,$$

$$\frac{(a_2-\rho_1)\cdots(a_2-\rho_6)}{\phi'(a_2)} + \frac{(b_2-\rho_1)\cdots(b_2-\rho_6)}{\phi'(b_2)} = 0,$$

$$\frac{(a_3-\rho_1)\cdots(a_3-\rho_6)}{\phi'(a_3)} + \frac{(b_3-\rho_1)\cdots(b_3-\rho_6)}{\phi'(b_3)} = 0,$$

$$\frac{(a_4-\rho_1)\cdots(a_4-\rho_6)}{\phi'(a_4)} + \frac{(b_4-\rho_1)\cdots(b_4-\rho_6)}{\phi'(b_4)} = 0.$$

From the theory of partial fractions the sum of the above expressions is identically zero and therefore the equations (2) amount to three equations.

3. *The tritangent circles.* If we consider the circle to be tritangent to the bicubic, the values of the parameters of the points of intersection become equal by pairs. Suppose $\rho_1 = \rho_4$, $\rho_2 = \rho_5$, $\rho_3 = \rho_6$, then (2) becomes,

$$\frac{(a_i-\rho_1)^2(a_i-\rho_2)^2(a_i-\rho_3)^2}{(b_i-\rho_1)^2(b_i-\rho_2)^2(b_i-\rho_3)^2} = k_i$$

and we have

$$(3) \qquad \begin{cases} \dfrac{(a_1-\rho_1)(a_1-\rho_2)(a_1-\rho_3)}{(b_1-\rho_1)(b_1-\rho_2)(b_1-\rho_3)} = (k_1)^{1/2}, \\[2ex] \dfrac{(a_2-\rho_1)(a_2-\rho_2)(a_2-\rho_3)}{(b_2-\rho_1)(b_2-\rho_2)(b_2-\rho_3)} = (k_2)^{1/2}, \\[2ex] \dfrac{(a_3-\rho_1)(a_3-\rho_2)(a_3-\rho_3)}{(b_3-\rho_1)(b_3-\rho_2)(b_3-\rho_3)} = (k_3)^{1/2}, \\[2ex] \dfrac{(a_4-\rho_1)(a_4-\rho_2)(a_4-\rho_3)}{(b_4-\rho_1)(b_4-\rho_2)(b_4-\rho_3)} = (k_4)^{1/2}. \end{cases}$$

Now we have,

$$k_1k_2k_3k_4 = \prod^6 (a_1 - a_2) / \prod^6 (b_1 - b_2)^2.$$

and therefore,

$$(k_1)^{1/2}(k_2)^{1/2}(k_3)^{1/2}(k_4)^{1/2}$$
$$= \pm \frac{(a_1-a_2)(a_1-a_3)(a_1-a_4)(a_2-a_3)(a_2-a_4)(a_3-a_4)}{(b_1-b_2)(b_1-b_3)(b_1-b_4)(b_2-b_3)(b_2-b_4)(b_3-b_4)};$$

or writing $| a_1\, a_2\, a_3\, a_4 |$ for the determinant

$$\begin{vmatrix} a_1^3 & a_1^2 & a_1 & 1 \\ a_2^3 & a_2^2 & a_2 & 1 \\ a_3^3 & a_3^2 & a_3 & 1 \\ a_4^3 & a_4^2 & a_4 & 1 \end{vmatrix}$$

we have

(4) $$(k_1)^{1/2}(k_2)^{1/2}(k_3)^{1/2}(k_4)^{1/2} = \pm\, | a_1\, a_2\, a_3\, a_4 | \,/\, | b_1\, b_2\, b_3\, b_4 |;$$

and the question is to determine the sign. If we eliminate from (3) the symmetric functions of ρ_i we have a determinant whose first row is,

$$a_1^3 - (k_1)^{1/2}b_1^3, \qquad a_1^2 - (k_1)^{1/2}b_1^2, \qquad a_1 - (k_1)^{1/2}b_1, \qquad 1 - (k_1)^{1/2}.$$

Expanded, this consists of pairs of determinants such as:

$$| a_1\, a_2\, a_3\, a_4 | + (k_1)^{1/2}(k_2)^{1/2}(k_3)^{1/2}(k_4)^{1/2} | b_1\, b_2\, b_3\, b_4 |$$
$$- (k_1)^{1/2} | b_1\, a_2\, a_3\, a_4 | - (k_2)^{1/2}(k_3)^{1/2}(k_4)^{1/2} | a_1\, b_2\, b_3\, b_4 |$$
$$+ (k_1)^{1/2}(k_2)^{1/2} | b_1\, b_2\, a_3\, a_4 | + (k_3)^{1/2}(k_4)^{1/2} | a_1\, a_2\, b_3\, b_4 |$$

and each pair vanishes when

(4′) $$(k_1)^{1/2}(k_2)^{1/2}(k_3)^{1/2}(k_4)^{1/2} = - | a_1\, a_2\, a_3\, a_4 | \,/\, | b_1\, b_2\, b_3\, b_4 |.$$

For, in this notation,

$$k_1 = \frac{| a_1\, a_2\, a_3\, a_4 |}{| b_1\, b_2\, b_3\, b_4 |} \cdot \frac{| a_1\, b_2\, b_3\, b_4 |}{| b_1\, a_2\, a_3\, a_4 |},$$

so that the second pair row vanishes, and

$$k_1k_2 = \frac{| a_1\, a_2\, a_3\, a_4 |}{| b_1\, b_2\, b_3\, b_4 |} \cdot \frac{| a_1\, a_2\, b_3\, b_4 |}{| b_1\, b_2\, a_3\, a_4 |},$$

so that the third pair row vanishes. We must then in (4) pick the negative sign. When we select $(k_1)^{1/2}$, $(k_2)^{1/2}$, and $(k_3)^{1/2}$, $(k_4)^{1/2}$ is now uniquely determined. Three of the equations (3) determine the fourth; and, since we can choose $(k_1)^{1/2}$, $(k_2)^{1/2}$, $(k_3)^{1/2}$ in 8 ways, there are 8 tritangent circles.

If we denote one of these by $+, +, +, +$ instead of by $(k_1)^{1/2}$, $(k_2)^{1/2}$, $(k_3)^{1/2}$, $(k_4)^{1/2}$, then there are, by equation (4), six like $+, +, -, -$, and the eighth is $-, -, -, -$.

4. *The biquadratics.* Consider now a biquadratic curve given by a self-conjugate equation of the second degree in x and in $\bar{x}$. Substituting for x and $\bar{x}$ from (1) we have an equation of degree 12 in ρ; so that the bicubic and the biquadratic have 12 common points.

As before, Abel's theorem takes the form

$$\frac{(a_i-\rho_1)\cdots(a_i-\rho_{12})}{(b_i-\rho_1)\cdots(b_i-\rho_{12})}=c_i.$$

To determine c_i we consider the circle on the other double points, taken twice. We obtain thus $c_i=k_i^2$.

Thus the four relations on the common points are

$$\frac{(a_i-\rho_1)\cdots(a_i-\rho_{12})}{(b_i-\rho_1)\cdots(b_i-\rho_{12})}=k_i^2. \tag{5}$$

These are independent relations.

If now we write the conditions for four contact circles

$$\frac{(a_i-\rho_1)(a_i-\rho_2)(a_i-\rho_3)}{(b_i-\rho_1)(b_i-\rho_2)(b_i-\rho_3)}=\pm(k_i)^{\frac{1}{2}},$$

$$\frac{(a_i-\rho_4)(a_i-\rho_5)(a_i-\rho_6)}{(b_i-\rho_4)(b_i-\rho_5)(b_i-\rho_6)}=\pm(k_i)^{\frac{1}{2}},$$

$$\frac{(a_i-\rho_7)(a_i-\rho_8)(a_i-\rho_9)}{(b_i-\rho_7)(b_i-\rho_8)(b_i-\rho_9)}=\pm(k_i)^{\frac{1}{2}},$$

$$\frac{(a_i-\rho_{10})(a_i-\rho_{11})(a_i-\rho_{12})}{(b_i-\rho_{10})(b_i-\rho_{11})(b_i-\rho_{12})}=\pm(k_i)^{\frac{1}{2}},$$

the relations (5) are satisfied, if we take an even number of negative signs.

Taking four such circles, for example

1)	+	+	+	+
2)	—	+	+	—
3)	+	—	+	—
4)	—	—	+	+

their 12 points of contact are on a biquadratic. Four such circles are called a set.

Any three circles have 9 points of contact on a biquadratic. For the above rule of signs determines the fourth circle. We have an even number of negative signs in each row by equation (4), so that if we have an even number in three columns we must have an even number in the fourth.

We have now 8 circles, and 56 ways of selecting three. With each three is a fourth. Hence there are 14 sets of four, that is 14 biquadratics.

5. *The extension of Hart's theorem.* Let us denote the form in x and $\bar{x}$ which gives the bicubic by C, that for one of the biquadratics by B, and that for the set of four circles by A_1, A_2, A_3, A_4. Then since B is on the points of contact of A_i and C, there is an identity

$$A_1A_2A_3A_4 - B^2 = CA, \tag{6}$$

where A is a bilinear form. Hence A_1 touches CA where it meets B. It already touches C three times; it therefore touches A at the fourth intersection of A_1 and B. This is also true for A_2, A_3, and A_4. Therefore the circles of a set are touched by a circle. Since there are fourteen distinct biquadratics, there are fourteen such circles.

6. *The general rational curve.* The argument of this paper applies to any rational curve on a sphere. We have now

$$x = (\alpha\rho)^n/(\beta\rho)^n$$

with $(n-1)^2$ double points. Let the parameters at these be a_i, b_i. Omitting one double point, there is on the $n(n-2)$ others a definite curve

$$f^{n-2} \equiv f(x^{n-2}, \bar{x}^{n-2}) = 0.$$

This does not meet the given curve elsewhere. For the general f^{n-2} Abel's theorem is

$$\prod_i {}_{2n(n-2)} \equiv \frac{(a_i-\rho_1)\cdots(a_i-\rho_{2n(n-2)})}{(b_i-\rho_1)\cdots(b_i-\rho_{2n(n-2)})} = k_i, \qquad [i = 1\cdots(n-1)^2].$$

As in § 2 these $(n-1)^2$ equations amount only to $n(n-2)$.

For a contact curve we have

$$\prod_i {}_{n(n-2)} \equiv \frac{(a_i-\rho_1)\cdots(a_i-\rho_{n(n-2)})}{(b_i-\rho_1)\cdots(b_i-\rho_{n(n-2)})} = (k_i)^{1/2}.$$

By the determinant argument of § 3, the product of these selected square roots is definitive. We can only change the sign of an even number, therefore the number of contact curves is, writing μ for $(n-1)^2$,

$$1 + \binom{\mu}{2} + \binom{\mu}{4} + \cdots,$$

that is $2^{\mu-1}$ or $2^{n(n-2)}$. They fall again into sets of 4 which lie on a $f^{2(n-2)}$. And the identity of § 5 now becomes

$$f_1^{n-2}f_2^{n-2}f_3^{n-2}f_4^{n-2} - [f^{2(n-2)}]^2 = f^n f^{3n-8}.$$

INTEGRAL EQUATIONS AND THE COOLING PROBLEM FOR SEVERAL MEDIA.

By W. M. Rust, Jr.

PART I. Solution of a Generalized Abel Integral Equation.

In a simple problem concerning the cooling of castings, the determination of the temperature may be reduced to the solution of a certain type of generalized Abel integral equation. The purpose of this first part is to discuss a method of solving this equation.

The equation can be put in the form

$$(1.1)\quad \int_0^t \left\{ \frac{1+K(t,t')}{(t-t')^{1/2}} \right\} u(t')\,dt' = f(t) \text{ with } K(t,t') \text{ and } f(t) \text{ known.}$$

In the actual problem, the function $K(t,t')$ vanishes to a high order for $t=t'$. For the method used here it is sufficient that $K(t,t')$ satisfy the conditions;

(A) $K(t,t')/(t-t')$ is bounded and is absolutely continuous in t, uniformly for all $t' \leq t$, and is summable in each of the variables. This implies the vanishing to at least the first order of $K(t,t')$ for $t=t'$.

(B) $\partial K(t,t')/\partial t'$ exists, is bounded and is absolutely continuous in t, uniformly for all $t' \leq t$, and is summable in each of the variables.*

The function $f(t)$ is assumed bounded and absolutely continuous, it is *not* assumed that $\operatorname*{Limit}_{t=0} f(t) = 0$, as is usually a condition for solution.

We shall assume that the equation (1.1) has a summable solution and

* The form (1.1) was used because this is the form needed in the sequel, however no essential change is made if we write the equation in the form

$$\int_0^t [G(t,t')/(t-t')^{\lambda}]\, u(t')\,dt' = f(t).$$

Where $G(t,t')$ satisfies the following conditions

(1) $G(t,t')$ is bounded and is absolutely continuous in t, uniformly for all $t' \leq t$, and is summable in each of the variables.

(2) For t' near enough to t, we have $G(t,t') = g(t) + h(t)(t-t')$ plus higher powers of $(t-t')$, and $|g(t)| \geq m > 0$, and $g(t)$ and $h(t)$ are bounded and absolutely continuous.

(3) $\partial G(t,t')/\partial t'$ exists, is bounded and is absolutely continuous in t, uniformly for all $t' \leq t$, and is summable in each of the variables, λ is a constant between zero and one, $f(t)$ is bounded and absolutely continuous.

shall solve for $\int_0^{t''} u(t)\,dt$, rather than for $u(t)$ itself. The method of solution is to multiply through by $(t''-t)^{-1/2}$ and integrate from $t=0$ to $t=t''$, as in the usual treatment. If (1.1) is satisfied by $u(t)$, this gives

$$(1.2)\qquad \int_0^{t''} \frac{dt}{(t''-t)^{1/2}} \int_0^{t} \left\{ \frac{1+K(t,t')}{(t-t')^{1/2}} \right\} u(t)\,dt' = \int_0^{t''} \frac{f(t)\,dt}{(t''-t)^{1/2}}.$$

In the left hand member we change the order of integration, which is allowable since $K(t,t')$ is bounded. The result is

$$\int_0^{t''} \frac{dt}{(t''-t)^{1/2}} \int_0^{t} \left\{ \frac{1+K(t,t')}{(t-t')^{1/2}} \right\} u(t')\,dt'$$
$$= \pi \int_0^{t''} u(t')\,dt' + \int_0^{t''} u(t')\,dt' \int_{t'}^{t''} \frac{K(t,t')\,dt}{(t''-t)^{1/2}(t-t')^{1/2}}.$$

At this point we vary from the usual treatment by integrating the last term by parts. This gives

$$\left[\left\{ \int_0^{t'} u(t)\,dt \right\} \left\{ \int_{t'}^{t''} \frac{K(t,t')\,dt}{(t''-t)^{1/2}(t-t')^{1/2}} \right\} \right]_{t'=0}^{t'=t''}$$
$$- \int_0^{t''} \left\{ \int_0^{t'} u(t)\,dt \right\} \frac{\partial}{\partial t'} \left(\int_{t'}^{t''} \frac{K(t,t')\,dt}{(t''-t)^{1/2}(t-t')^{1/2}} \right) dt'.$$

The term outside the sign of integration vanishes since the factor $\int_0^{t'} u(t)\,dt$ vanishes for $t'=0$ and the factor $\int_{t'}^{t''} \frac{K(t,t')\,dt}{(t''-t)^{1/2}(t-t')^{1/2}}$ vanishes for $t'=t''$. This follows from condition (A), for if $K(t,t')$ vanishes to the order α for $t=t'$, this integral is less than

$$\int_{t'}^{t''} \frac{M\,dt}{(t''-t)^{1/2}(t-t')^{1/2-\alpha}} = M(t''-t')^{\alpha} B(1/2, 1/2+\alpha)$$

and vanishes to the order α for $t'=t''$.

Formal differentiation gives

$$(1.3)\qquad \frac{\partial}{\partial t'} \left(\int_{t'}^{t''} \frac{K(t,t')\,dt}{(t''-t)^{1/2}(t-t')^{1/2}} \right) = \left[\frac{-K(t,t')}{(t''-t)^{1/2}(t-t')^{1/2}} \right]_{t=t''}$$
$$+ \int_{t'}^{t''} \left\{ \frac{(\partial/\partial t')K(t,t')}{(t''-t)^{1/2}(t-t')^{1/2}} + \frac{K(t,t')}{2(t''-t)^{1/2}(t-t')^{3/2}} \right\} dt.$$

The first term is zero by condition (A). The second term by condition (B) is bounded. The third term is bounded, for by condition (A) the function $K(t,t')$ vanishes to at least the first order for $t=t'$ and so the integral is less than

$$M\int_{t'}^{t''}(t''-t)^{-1/2}(t-t')^{-3/2+\alpha}dt$$

with α at least as great as one.

This establishes the fact that the derivative (1.3) is bounded and justifies the integration by parts.

We thus have as a necessary condition on solutions of (1.1),

$$(1.4)\qquad \pi\int_0^{t''}u(t')\,dt' \equiv \int_0^{t''}\frac{f(t')\,dt'}{(t''-t')^{1/2}} + \int_0^{t''}\left\{\int_0^{t'}u(t)\,dt\right\}\frac{\partial}{\partial t'}\left(\int_{t'}^{t''}\frac{K(t,t')\,dt}{(t''-t)^{1/2}(t-t')^{1/2}}\right)dt'.$$

We wish to show that the equation

$$(1.5)\qquad \pi V(t'') = \int_0^{t''}\frac{f(t')\,dt'}{(t''-t')^{1/2}} + \int_0^{t''}V(t')\,\frac{\partial}{\partial t'}\left(\int_{t'}^{t''}\frac{K(t,t')\,dt}{(t''-t)^{1/2}(t-t')^{1/2}}\right)dt'$$

has a unique absolutely continuous solution, whose derivative satisfies the equation (1.1) nearly everywhere.

Since equation (1.5) is a Volterra integral equation of the second sort with a bounded kernel, (1.3), it may be solved by a process of successive approximations, the process necessarily converging and the solution being summable and unique.*

Since the equation (1.4) is a necessary condition on summable solutions of (1.1), the uniqueness of the solution of (1.5) shows that the solutions of (1.1) can, at most, differ on a point set of zero measure.

To show that the solution of (1.5) is absolutely continuous, we need a lemma.

LEMMA I. *The function* $F(t) = \int_0^1 f(t,t')g(t')\,dt'$ *is an absolutely continuous function of* t *if* $f(t,t')$ *is absolutely continuous in* t, *uniformly for all* t' *and if* $g(t')$ *is a summable function.*

Proof. By the λ-variation of the function $f(t,t')$ we shall mean the function $\tau_f(\lambda)$ defined as the upper limit, for all sets of non-overlapping intervals (a_i, b_i) such that $\Sigma\,|a_i - b_i| \leq \lambda$ and for all t', of the numbers

$$\Sigma\,|f(a_i,t') - f(b_i,t')|.$$

---

* Volterra, *Leçons sur les Équations Intégrales*, p. 40 ff.

To say that $f(t, t')$ is absolutely continuous in t, uniformly for all t', means that the λ-variation approaches zero with λ.

The λ-variation of $F(t)$, defined in an analogous way, is less than or equal to

$$\tau_f(\lambda) \int_0^1 g(t')\,dt'.$$

This approaches zero with λ, so that $F(t)$ is absolutely continuous. This establishes the Lemma.*

By means of the transformation $t' = t''y$, the first term on the right hand side of equation (1.5) becomes

$$(t'')^{1/2} \int_0^1 f(t''y)(1-y)^{-1/2}\,dy$$

which satisfies the condition of the Lemma I and so is absolutely continuous.

By means of the transformation $t = t' + y(t'' - t')$, the expression (1.3) becomes

$$\int_0^1 \left\{ \frac{K_2(t' + y\,[t'' - t'], t')}{(1-y)^{1/2}\,y^{1/2}} + \frac{K(t' + y\,[t'' - t'], t')}{2(1-y)^{1/2}y^{3/2}(t'' - t')} \right\} dt$$

where $K_2(t, t') = \partial K(t, t')/\partial t'$. By condition (B) and the Lemma I, the first term is absolutely continuous in t''. By condition (A) $K/(t - t') = K/y(t'' - t')$ is absolutely continuous in t'', and bounded so that by the Lemma I the second term is absolutely continuous in t''. In each case the absolute continuity is uniform in t'. Thus the expression (1.3) is absolutely continuous in t'', uniformly in t'.

We need now the further Lemma,

LEMMA II. *The function* $G(t'') = \int_0^{t''} g(t')k(t'', t')\,dt'$ *is an absolutely continuous function of* t'', *if* $k(t'', t')$ *is absolutely continuous in* t'' *uniformly for all* t' *and bounded, and if* $g(t')$ *is summable.*

Proof. The λ-variation, $\tau_G(\lambda)$, of $G(t'')$ is the upper limit for all sets of non-overlapping intervals (a_i, b_i) such that $\Sigma\,|\,a_i - b_i\,| \leq \lambda$ of the numbers

$$\Sigma \left| \int_0^{b_i} g(t')k(b_i, t')\,dt' - \int_0^{a_i} g(t')k(a_i, t')\,dt' \right|$$

$$= \Sigma \left| \int_{a_i}^{b_i} g(t')k(b_i, t')\,dt' + \int_0^{a_i} g(t')\{k(b_i, t') - k(a_i, t')\}\,dt' \right|.$$

* The condition is a sufficient one, regardless of the manner in which $f(t, t')$ involves t, but in general it is not a necessary condition. If $f(t, t')$ is a function of t alone, the condition is obviously necessary, but if $f(t, t')$ is a function of the product tt', the condition may be modified to admit a set of points of discontinuity of the first sort, if this set is of zero measure.

Thus

$$\tau_G(\lambda) \leq \overline{\text{Limit}}\, \Sigma\, M \int_{a_i}^{b_i} |\, g(t')\,|\, dt' + \tau_k(\lambda) \int_0^{T''} |\, g(t')k \,|\, t',$$

M is maximum absolute value of $k(t,t')$,

with T'' the greatest of the numbers a_i or b_i. Since $g(t')$ is summable, each of these terms approaches zero with λ. This establishes the Lemma.

The second term in the right-hand member of (1.5) satisfies the conditions of Lemma II and so is absolutely continuous. Since each of the terms in the right-hand member of (1.5) is absolutely continuous, $V(t'')$ is absolutely continuous as was to be shown.

Since $V(t)$ is absolutely continuous, it has a derivative $V'(t)$ nearly everywhere and $\int_0^t V'(t')dt' = V(t)$. We shall now show that $u(t) = V'(t)$ satisfies equation (1.1) nearly everywhere. To do this we write

$$D(t) = \int_0^t \left\{ \frac{1+K(t,t')}{(t-t')^{1/2}} V'(t') \right\} dt' - f(t). \tag{1.6}$$

Multiplying through by $(t''-t)^{-1/2}$, integrating from $t=0$ to $t=t''$, changing the order of integration and integrating by parts, as above, gives

$$\begin{aligned}\int_0^{t''} \frac{D(t)\,dt}{(t''-t)^{1/2}} &\equiv \pi \int_0^{t''} V'(t)\,dt - \int_0^{t''} \frac{f(t)\,dt}{(t''-t)^{1/2}} \\ &- \int_0^{t''} \left\{ \int_0^{t'} V'(t)\,dt \right\} \frac{\partial}{\partial t'} \left(\int_{t'}^{t''} \frac{K(t,t')\,dt}{(t''-t)^{1/2}(t-t')^{1/2}} \right) dt' \equiv \pi V(t'') \\ &- \int_0^{t''} \frac{f(t)\,dt}{(t''-t)^{1/2}} - \int_0^{t''} V(t') \frac{\partial}{\partial t'} \left(\int_{t'}^{t''} \frac{K(t,t')\,dt}{(t''-t)^{1/2}(t-t')^{1/2}} \right) dt'.\end{aligned} \tag{1.7}$$

But since $V(t)$ satisfies equation (1.5), the right-hand side of (1.7) is zero. Hence, as Tonelli * shows, $D(t)$ is zero nearly everywhere, that is, the derivative of the solution of equation (1.5) satisfies (1.1) nearly everywhere.

We have the following theorem.

THEOREM. *Under the conditions specified with respect to* $K(t,t')$ *and* $f(t)$ *in (A) and (B), the equation* (1.1) *has one and only one summable solution* $u(t)$, *and this solution is given as the derivative of the solution of equation* (1.5).

Two summable functions differing only on a point set of zero measure are, of course, to be considered equivalent.

* L. Tonelli, "Su un problema di Abel," *Mathematische Annalen*, Vol. 99 (1928), p. 187.

PART II. A PROBLEM IN THE FLOW OF HEAT.

As an example of a physical problem that may be solved by such an equation, we consider the following problem in heat conduction, one which has both theoretical and intrinsic interest.

A quantity of one material is heated and placed between two masses of a different material, as a casting in its mold. The problem is: given the initial temperatures of the three regions and the temperatures at the two outer boundaries, find the temperature at any point at any time.

For convenience we consider very simple conditions. We take the three regions to be bounded by parallel infinite planes and take the thicknesses of the two outer regions to be equal. We take the initial temperatures in the outer regions to be constant and equal to each other. We take the initial temperature of the inner region to be constant, not necessarily the same as in the outer regions. We take the temperatures on the outer bounding surfaces to be equal and, at any instant, constant over the entire bounding plane. With these conditions the problem is symmetric about the central plane and at any instant the temperature is constant over any plane parallel to the central plane, thus only one spatial coördinate, the distance from the central plane, is involved. Some generalizations of this problem are considered in Part V.

The conductivity of the material in the inner region is K_1 and the conductivity of the material in the outer regions is K_2. The quantities a^2 and b^2 are positive constants, equal to the ratios of the conductivity to the product of the specific heat by the density, for the inner and outer regions respectively. Also, x is the distance from the central plane and t is the time after the initial time.

We take the bounding planes between the inner and outer regions to be at $x = m$ and $x = -m$ and the outer boundaries to be at $x = l$ and $x = -l$.

At interior points, $u_1(x, t)$, the temperature in the inner region, and $u_2(x, t)$, the temperature in the outer regions, satisfy the partial differential equations

$$(2.1) \quad \begin{cases} \partial^2 u_1(x,t)/\partial x^2 - a^2(\partial u_1(x,t)/\partial t) = 0 & -m < x < m \\ \partial^2 u_2(x,t)/\partial x^2 - b^2(\partial u_2(x,t)/\partial t) = 0 & -l < x < -m \text{ or } m < x < l \end{cases}$$

respectively.*

* For the derivation of the equations (2.1) see any standard work on the conduction of heat, for example Carslaw, *Conduction of Heat*, Chapter One, or Riemann-Weber, *Differentialgleichungen der Physik*, Vol. 2 (1912), p. 82.

13

If the temperature in the inner region is initially u_1, a constant, and in the outer regions u_2, a constant, we have

$$(2.2)\quad \begin{cases} \operatorname*{Limit}_{t=0+} u_1(x,t) = u_1 & \text{for} \quad -m < x < m \\ \operatorname*{Limit}_{t=0+} u_2(x,t) = u_2 & \text{for} \quad -l < x < -m \quad \text{or} \quad m < x < l. \end{cases}$$

At the outer boundaries the temperature is taken to be a known, bounded, continuous function of the time, say $f(t)$, with a bounded derivative, that is

$$(2.3)\qquad \operatorname*{Limit}_{x=l-0} u_2(x,t) = f(t) \quad \text{for} \quad t > 0.$$

At the separating boundaries we have two conditions, first, the temperature is continuous in x across the boundaries, that is,

$$(2.4)\qquad \operatorname*{Limit}_{x=m-0} u_1(x,t) = \operatorname*{Limit}_{x=m+0} u_2(x,t) \quad \text{for} \quad t > 0,$$

and, second, the partial derivatives with respect to x satisfy the equation †

$$(2.5)\quad \operatorname*{Limit}_{x=m-0} K_1 \partial u_1(x,t)/\partial x = \operatorname*{Limit}_{x=m+0} K_2 \partial u_2(x,t)/\partial x \quad \text{for} \quad t > 0.$$

Similar conditions are imposed at $x = -m$ and at $x = -l$, but due to the symmetry of the problem, these are equivalent to those given by the equations (2.3) to (2.5).

We establish the following uniqueness theorem.

UNIQUENESS THEOREM A. *There can not be more than one solution of the problem as given by equations* (2.1), *subject to the conditions* (2.2) *to* (2.5), *which is bounded everywhere* (*including* $t = 0$), *is continuous for* $t > 0$, *except at the boundaries, and has first derivatives with respect to each of the variables, which are continuous for* $t > 0$, *except at the boundaries—the derivative with respect to* x *being bounded everywhere for* $t > 0$.

Suppose there were two such solutions. Call the difference between these solutions $V_1(x,t)$ and $V_2(x,t)$ in the inner and outer regions respectively. $V_1(x,t)$ and $V_2(x,t)$ satisfy the conditions of the theorem and take on the boundary and initial values zero.

Consider the integrals

$$\begin{aligned} (\mathrm{a})\qquad J_\epsilon(t) = (K_2 b^2/2) &\int_{-l+\epsilon}^{-m-\epsilon} (V_2(x,t))^2\,dx \\ &+ (K_1 a^2/2) \int_{-m+\epsilon}^{m-\epsilon} (V_1(x,t))^2\,dx \\ &+ (K_2 b^2/2) \int_{m+\epsilon}^{l-\epsilon} (V_2(x,t))^2\,dx. \end{aligned}$$

† For the derivation of equations (2.5) see, for example, Riemann-Weber, Vol. 2 (1912), p. 85, cited above.

For $t > 0$, $V_i(x, t)$ and $\partial V_i(x, t)/\partial t$ are continuous and bounded in the region of integration, so that we may differentiate under the integral sign. We have

$$\frac{dJ_\epsilon(t)}{dt} = K_2 b^2 \int_{-l+\epsilon}^{-m-\epsilon} V_2(x, t) \frac{\partial V_2}{\partial t}(x, t)\,dx$$

$$+ K_1 a^2 \int_{-m+\epsilon}^{m-\epsilon} V_1(x, t) \frac{\partial V_1}{\partial t}(x, t)\,dx + K_2 b^2 \int_{m+\epsilon}^{l-\epsilon} V_2(x, t) \frac{\partial V_2}{\partial t}(x, t)\,dx$$

and

(b) $$\int_{t_1}^{t_2} \frac{dJ_\epsilon(t)}{dt}\,dt = J_\epsilon(t_2) - J_\epsilon(t_1), \qquad t_1, t_2, > 0$$

applying (2.1), we have

$$\frac{dJ_\epsilon(t)}{dt} = K_2 \int_{-l+\epsilon}^{-m-\epsilon} V_2(x, t) \frac{\partial^2 V_2}{\partial x^2}(x, t)\,dx$$

$$+ K_1 \int_{-m-\epsilon}^{m-\epsilon} V_1(x, t) \frac{\partial^2 V_1}{\partial x^2}(x, t)\,dx + K_2 \int_{m+\epsilon}^{l-\epsilon} V_2(x, t) \frac{\partial^2 V_2}{\partial x^2}(x, t)\,dx.$$

In the regions of integration each $\partial^2 V_i/\partial x^2$ is continuous and bounded, so that we may integrate by parts.

(c) $$dJ_\epsilon/dt = K_2 V_2(x, t)\,(\partial V_2/\partial x)(x, t)\Big]_{x=-l+\epsilon}^{x=-m-\epsilon}$$

$$+ K_1 V_1(x, t)\,(\partial V_1/\partial x)(x, t)\Big]_{x=-m-\epsilon}^{x=+m-\epsilon} + K_2 V_2(x, t)\,(\partial V_2/\partial x)(x, t)\Big]_{x=m+\epsilon}^{x=l-\epsilon}$$

$$- K_2 \int_{-l+\epsilon}^{-m-\epsilon} ((\partial V_2/\partial x)(x, t))^2\,dx - K_1 \int_{-m+\epsilon}^{+m-\epsilon} ((\partial V_1/\partial x)(x, t))^2\,dx$$

$$- K_2 \int_{m+\epsilon}^{l-\epsilon} ((\partial V_2/\partial x)(x, t))^2\,dx.$$

We have then from (b) and (c),

(d) $$\operatorname*{Limit}_{\epsilon=0}\,[J_\epsilon(t_2) - J_\epsilon(t_1)] = \operatorname*{Limit}_{\epsilon=0} \int_{t_1}^{t_2} (dJ_\epsilon/dt)\,dt$$

$$= \operatorname*{Limit}_{\epsilon=0} \int_{t_1}^{t_2} \Big\{ K_2 V_2(x, t)\,(\partial V_2/\partial x)(x, t)\Big]_{x=-l+\epsilon}^{x=-m-\epsilon}$$

$$+ K_1 V_1(x, t)\,(\partial V_1/\partial x)(x, t)\Big]_{x=-m+\epsilon}^{x=+m-\epsilon} + K_2 V_2(x, t)\,(\partial V_2/\partial x)(x, t)\Big]_{x=m+\epsilon}^{x=l-\epsilon}$$

$$- K_2 \int_{-l+\epsilon}^{-m-\epsilon} ((\partial V_2/\partial x)(x, t))^2\,dx - K_1 \int_{-m+\epsilon}^{+m-\epsilon} ((\partial V_1/\partial x)(x, t))^2\,dx$$

$$- K_2 \int_{m+\epsilon}^{l-\epsilon} ((\partial V_2/\partial x)(x, t))^2\,dx \Big\}\,dt.$$

Since each $V_i(x, t)$ is bounded

$$\text{(e)} \qquad \operatorname*{Limit}_{\epsilon=0} J_\epsilon(t) = J(t) = K_2 b^2/2 \int_{-l}^{-m} (V_2(x,t))^2\, dx$$
$$+ K_1 a^2/2 \int_{-m}^{+m} (V_1(x,t))^2\, dx + K_2 b^2/2 \int_{m}^{l} (V_2(x,t))^2\, dx.$$

The left-hand member of (d) is thus

$$J(t_2) - J(t_1).$$

In the right-hand member of (d), we may pass to the limit under the sign of integration in the first three terms, since $V_i(x,t)$ and $\partial V_i/\partial x(x,t)$ are bounded.* That is

$$\text{(f)} \qquad \operatorname*{Limit}_{x=x_0\pm 0} \int_{t_1}^{t_2} K_i V_i(x,t)\,(\partial V_i/\partial x)\,(x,t)\, dt$$
$$= \int_{t_1}^{t_2} \operatorname*{Limit}_{x=x_0\pm 0} \{K_i V_i(x,t)\,(\partial V_i/\partial x)\,(x,t)\}\, dt$$

x_0 being the x coördinate of any boundary. But by the boundary conditions the limit of the first three terms is zero.†

Since K_1 and K_2 are positive constants, the remaining three terms are negative, or zero, if $t_2 > t_1$, so that in the limit the right-hand member of (d) is negative or zero.

We have then

$$\text{(g)} \qquad J(t_2) - J(t_1) \leq 0 \quad \text{if} \quad t_2 \geq t_1 > 0.$$

Since $[V_i(x,t)]^2$ is bounded and $\operatorname*{Limit}_{t=0} V_i(x,t) = 0$, nearly everywhere, we have

$$\operatorname*{Limit}_{t=0} J(t) = 0.$$

Applying this to (g) we have

$$\text{(h)} \qquad J(t) \leq 0, \quad \text{if} \quad t > 0.$$

But in (e) the terms are all positive, since K_1 and K_2 are positive, so that

$$\text{(j)} \qquad J(t) \geq 0, \quad \text{if} \quad t > 0.$$

We see (h) and (j) can only be satisfied simultaneously if

$$J(t) = 0.$$

---

* See Lebesgue, *Leçons sur L'integration* (1928), p. 125.

† For infinite outer regions, we may replace (2.3) by the condition

$$\operatorname{Limit} \; (\partial u/\partial x)\,(x,t) \equiv 0, \; t > 0.$$

But from (e) this can only be the case if $V_i(x,t)$ is zero nearly everywhere and since $V_i(x,t)$ is continuous except at the boundaries for $t > 0$, it must be identically zero for $t > 0$ and the two solutions are identical for $t > 0$. *This establishes the theorem.*

We now establish a uniqueness theorem with less restrictive conditions.

UNIQUENESS THEOREM B. *There can not be more than one solution of the problem satisfying the conditions of Theorem A, except that the derivative with respect to x is not assumed bounded, but summable in t and satisfying the condition*

$$\operatorname*{Limit}_{x=x_0\pm 0}\int_{t_1}^{t_2} |(\partial u_i/\partial x)(x,t)|\,dt = \int_{t_1}^{t_2}\operatorname*{Limit}_{x=x_0\pm 0} |(\partial u_i/\partial x)(x,t)|\,dt,$$

*x_0 being the x-coördinate of any boundary.**

The proof of this theorem is the same as that of Theorem A except in the proof of the equation (f)

$$\operatorname*{Limit}_{x=x_0\pm 0}\int_{t_1}^{t_2} K_iV_i(x,t)(\partial V_i/\partial x)(x,t)\,dt$$
$$= \int_{t_1}^{t_2}\operatorname*{Limit}_{x=x_0\pm 0}\{K_iV_i(x,t)(\partial V_i/\partial x)(x,t)\}\,dt.$$

In Theorem A this was a consequence of the boundedness of $V_i(x,t)$ and $\partial V_i(x,t)/\partial x$. Having removed this latter restriction, we need a different proof.

Each $V_i(x,t)$ is bounded, say less than M_i, hence

$$|K_iV_i(x,t)(\partial V_i/\partial x)(x,t)| \le K_iM_i\,|(\partial V_i/\partial x)(x,t)|$$

and by the condition of our theorem

$$\operatorname*{Limit}_{x=x_0\pm 0}\int_{t_1}^{t_2} |(\partial V_i/\partial x)(x,t)|\,dt = \int_{t_1}^{t_2}\operatorname*{Limit}_{x=x_0\pm 0} |(\partial V_i/\partial x)(x,t)|\,dt$$

and so †

---

* That is to say, the absolute continuity of $\int_0^t |(\partial u_i/\partial x)(x,t)|\,dt$—and hence of $\int_0^t (\partial u_i/\partial x)(x,t)\,dt$—is uniform in x. From this it follows that

$$\operatorname*{Limit}_{x=x_0\pm 0}\int_{t_1}^{t_2}(\partial u_i/\partial x)(x,t)\,dt = \int_{t_1}^{t_2}\operatorname*{Limit}_{x=x_0\pm 0}(\partial u_i/\partial x)(x,t)\,dt,$$

if $\operatorname*{limit}_{x=x_0\pm 0}(\partial u_i/\partial x)(x,t)$ exists.

† See Hobson, *Theory of Functions of a Real Variable*, Vol. 2 (1926), p. 290.

$$\operatorname*{Limit}_{x=x_0\pm 0}\int_{t_1}^{t_2} K_i V_i(x,t)\,(\partial V_i/\partial x)\,(x,t)\,dt$$
$$=\int_{t_1}^{t_2}\operatorname*{Limit}_{x=x_0\pm 0}\{K_i V_i(x,t)\,(\partial V_i/\partial x)\,(x,t)\}\,dt.$$

This establishes the equation (f) and the remainder of the proof is unaltered.

We will now show that a solution of the problem satisfying the conditions of the uniqueness theorem B *is given by*

$$(2.6)\qquad u_1(x,t)=u_1+\int_0^t\{(t-t')^{-\frac{1}{2}}e^{-a^2(x-m)^2/4(t-t')}+(t-t')^{-\frac{1}{2}}e^{-a^2(x+m)^2/4(t-t')}\}\,\psi_1(t')\,dt',$$

for $-m<x<m$,
and

$$(2.7)\qquad u_2(x,t)=u_2+\int_0^t(t-t')^{-\frac{1}{2}}e^{-b^2(x-m)^2/4(t-t')}\psi_2(t')\,dt'+\int_0^t(t-t')^{-\frac{1}{2}}e^{-b^2(x-l)^2/4(t-t')}\psi_3(t')\,dt',$$

for $m<x<l$, $[u_2(x,t)=u_2(-x,t)$ for $-l<x<m]$, *where* $\psi_1(t)$, $\psi_2(t)$ *and* $\psi_3(t)$ *are summable functions satisfying the following integral equations, almost everywhere.*

$$(2.8)\qquad \int_0^t(t-t')^{-\frac{1}{2}}\psi_3(t')\,dt'=f_2(t)-\int_0^t(t-t')^{-\frac{1}{2}}e^{-\beta^2/(t-t')}\psi_2(t')\,dt'.$$

$$(2.9)\qquad \psi_2(t)=-c\psi_1(t)+c\alpha\pi^{-\frac{1}{2}}\int_0^t(t-t')^{-3/2}e^{-\alpha^2/(t-t')}\psi_1(t')\,dt'+\beta\pi^{-\frac{1}{2}}\int_0^t(t-t')^{-3/2}e^{-\beta^2/(t-t')}\psi_3(t')\,dt'.$$

$$(2.10)\qquad \int_0^t\{(t-t')^{-\frac{1}{2}}e^{-\alpha^2/(t-t')}+(t-t')^{-\frac{1}{2}}\}\,\psi_1(t')\,dt'=f_1+\int_0^t(t-t')^{-\frac{1}{2}}\psi_2(t')\,dt'+\int_0^t(t-t')^{-\frac{1}{2}}e^{-\beta^2/(t-t')}\psi_3(t')\,dt'.$$

Here the integrals are taken to be Lebesgue integrals and we have used the following abbreviations:

(i) $f_2(t)=f(t)-u_2$
(ii) $\alpha=am$
(iii) $\beta=b(l-m)/2$
(iv) $c=K_1a/K_2b>0$
(v) $f_1=u_2-u_1$

By actual differentiation we see that

$$(t-t')^{-\frac{1}{2}} e^{-a^2(x-x')/4(t-t')}$$

when considered, as a function of x and t is a solution of the first equation (2.1) for any $x' \neq x$ and for any $t' < t$. If we replace the a^2 in the exponent by b^2, we have a solution of the second equation (2.1).*

The functions $u_1(x,t)$ and $u_2(x,t)$ as given by equations (2.6) and (2.7) are linear combinations of such solutions and except at the boundaries the integrals converge for any summable functions $\psi_1(t)$, $\psi_2(t)$ and $\psi_3(t)$. These expressions are easily seen to be continuous except at the boundaries. Likewise, if we form the first derivatives we get expressions which converge and are continuous except at the boundaries. The functions $u_1(x,t)$ and $u_2(x,t)$ thus satisfy the equation (2.1) in the inner and outer regions respectively.

In each of the integrals the first factor is bounded except at the boundaries so that each integrand is summable and the limit of each integral, as t approaches zero, is zero except at the boundaries. The initial conditions (2.2) are thus satisfied.

As a consequence of equations (2.8) to (2.10) we can show that $\psi_1(t)$, $\psi_2(t)$ and $\psi_3(t)$ are equivalent to functions of the form

$$A_i/t^{\frac{1}{2}} + \phi_i(t)$$

where $\phi_i(t)$ is bounded and continuous.

Equation (2.8) can be written

$$\text{(A)} \qquad \int_0^t (t-t')^{-\frac{1}{2}} \psi_3(t')\,dt' = g(t)$$

where

$$g(t) = f_2(t) - \int_0^t (t-t')^{-\frac{1}{2}} e^{-\beta^2/(t-t')} \psi_2(t')\,dt'.$$

In virtue of the conditions on $f(t)$ and the boundedness and continuity of $(t-t')^{-\frac{1}{2}} e^{-\beta^2/(t-t')}$ and the summability of $\psi_2(t)$, $g(t)$ is bounded and continuous. Its derivative

$$g'(t') = f'_1(t) - \int_0^t (\partial/\partial t)\{(t-t')^{-\frac{1}{2}} e^{-\beta^2/(t-t')}\}\,\psi_2(t')\,dt'$$

is likewise bounded and continuous.

Equation (A) is an Abel's equation of the usual type and has the solution †

* For a full discussion of the properties of these and similar functions, see E. E. Levi, "Sull' equazione del calore," *Annali di Matematica* (3), Vol. 14 (1908), p. 187.

† Volterra, *loc. cit.*, p. 37.

$$\psi_3(t) = g(0)/\pi t^{1/2} + 1/\pi \int_0^t (t-t')^{-1/2} g'(t')\,dt'.$$

Since $g'(t)$ is bounded and continuous, the same is true of

$$\int_0^t (t-t')^{-1/2} g'(t')\,dt'$$

and $\psi_3(t')$ has the required form.

To show that $\psi_2(t)$ and $\psi_1(t)$ have this form, we eliminate $\psi_2(t)$ from equations (2.9) and (2.10). Thus from (2.9)

(B) $$\psi_2(t) = -c\psi_1(t) + h(t)$$

where

$$h(t) = c\alpha\pi^{-1/2} \int_0^t (t-t')^{-3/2} e^{-\alpha^2/(t-t')}\psi_1(t')\,dt' + \beta\pi^{-1/2} \int_0^t (t-t')^{-3/2} e^{-\beta^2/(t-t')}\psi_3(t')\,dt'$$

and since $\psi_1(t)$ and $\psi_3(t)$ are summable, it is bounded and continuous and has a bounded and continuous derivative. We have then

$$\int_0^t (t-t')^{-1/2}\psi_2(t')\,dt' = -c \int_0^t (t-t')^{-1/2}\psi_1(t')\,dt' + \int_0^t (t-t')^{-1/2}h(t')\,dt' = -c \int_0^t (t-t')^{-1/2}\psi_1(t')\,dt' + c\int_0^t (t-t')^{-1/2} e^{-\alpha^2/(t-t')}\psi_1(t')\,dt + \int_0^t (t-t)^{-1/2} e^{-\beta^2/(t-t')}\psi_3(t')\,dt',$$

so that (2.10) gives

(C) $$\int_0^t (t-t')^{-1/2}\psi_1(t')\,dt' = h_1(t),$$

where

$$h_1(t) = f_1/(1+c) + 2/(1+c)\int_0^t (t-t')^{-1/2} e^{-\beta^2/(t-t')}\psi_3(t')\,dt' + (c-1)/(1+c)\int_0^t (t-t')^{-1/2} e^{-\alpha^2/(t-t')}\psi_1(t')\,dt',$$

and is bounded and continuous with a bounded and continuous derivative. So that, as above, equation (C) has the solution

$$\psi_1(t) = h_1(0)/\pi t^{1/2} + 1/\pi \int_0^t (t-t')^{-1/2} h'_1(t')\,dt,$$

and since $h_1'(t)$ is bounded and continuous, $\int_0^t (t-t')^{-1/2} h'_1(t')\,dt'$ is bounded and continuous and $\psi_1(t)$ has the required form.

From equation (B) is follows then that $\psi_3(t)$ also has this form.

Consider now a term of the form

$$\int_0^t (t-t')^{-1/2}\, e^{-a^2/(x-m)^2/4(t-t')}\psi_1(t')\,dt'.$$

In absolute value this is less than

$$\int_0^t (t-t')^{-1/2}\,|\,\psi_1(t')\,|\,dt'$$
$$\leq \int_0^t \pi^{-1}(t-t')^{-1/2}(t')^{-1/2}\,|\,h_1(0)\,|\,dt' + \int_0^t (t-t')^{-1/2}\,\phi_1(t')\,dt'$$

and so is bounded. Hence $u_1(x,t)$ and $u_2(x;t)$, being the sum of such terms and a finite constant, are bounded.

We also have

$$\operatorname*{Limit}_{x=l-0} \int_0^t (t-t')^{-1/2}\, e^{-b^2(x-l)^2/4(t-t')}\psi_3(t')\,dt' = \int_0^t (t-t')^{-1/2}\psi_3(t')\,dt'$$

since the integrand is less than a summable function, independent of x. We have then

$$\operatorname*{Limit}_{x=l-0}\ u_2(x,t) = u_2$$
$$+\int_0^t (t-t')^{-1/2}\, e^{-\beta^2/(t-t')}\psi_2(t')\,dt' + \int_0^t (t-t')^{-1/2}\,\psi_3(t')\,dt' = f(t)$$

in virtue of equation (2. 8). Thus the boundary condition (2. 3) is satisfied.

Now in the same way we can show

$$\operatorname*{Limit}_{x=m+0}\ u_2(x,t) = u_2$$
$$+\int_0^t (t-t')^{-1/2}\,\psi_2(t')\,dt' + \int_0^t (t-t')^{-1/2}\, e^{-\beta^2/(t-t')}\psi_3(t')\,dt'$$

and

$$\operatorname*{Limit}_{x=m-0}\ u_1(x,t) = u_1 + \int_0^t \{(t-t')^{-1/2}\, e^{-a^2/(t-t')} + (t-t')^{-1/2}\}\,\psi_1(t')\,dt',$$

so that from equation (2. 10), recalling (v), we have

$$\operatorname*{Limit}_{x=m-0}\ u_1(x,t) = \operatorname*{Limit}_{x=m+0}\ u_2(x,t)$$

and the boundary condition (2. 4) is satisfied.

To show that the boundary condition (2. 5) is satisfied we need to know the value of

$$\operatorname*{Limit}_{x=m+0}\ (\partial u_2/\partial x)\,(x,t).$$

By direct differentiation we have

$$(\partial u_1/\partial x)\,(x,t) = -a^2/2\int_0^t \{(t-t')^{-3/2}(x+m)\,e^{-a^2(x+m)^2/4(t-t')}$$
$$+ (t-t')^{-3/2}(x-m)\ e^{-a^2(x-m)^2/4(t-t')}\}\,\psi_1(t')\,dt',$$

and

$$(\partial u_2/\partial x)(x,t) = -b^2/2\int_0^t (t-t')^{-3/2}(x-m)e^{-b^2(x-m)^2/4(t-t')}\psi_2(t')\,dt'$$
$$-b^2/2\int_0^t (t-t')^{-3/2}(x-l)e^{-b^2(x-l)^2/4(t-t')}\psi_3(t')\,dt'.$$

The integrals involved all converge, except at the boundaries. To evaluate $\operatorname*{Limit}_{x=m+0} (\partial u_2/\partial x)(x,t)$, we need to know the value of

$$\operatorname*{Limit}_{x=m+0} b/2\pi^{1/2}\int_0^t (t-t')^{-3/2}(x-m)e^{-b^2(x-m)^2/4(t-t')}\phi(t')\,dt',$$

for $\phi(t)$ summable. Hobson * considers a more general problem. Restating his results in terms applicable here, we have the Theorem:

Let $F(t',t,x)$ be defined for $t',t>0$ and $m<x<l$. Let μ denote a positive number and let $F(t',t,x)$ satisfy the following conditions;

(1) For each pair of values t and x, and for all values of $t'>0$, such that $|t-t'|\geq\mu$, the function $F(t',t,x)$ is equivalent to a function that does not exceed in absolute value a positive number K_μ, independent of the values of t and x.

(2) If α and β are two numbers such that

$$0\leq\alpha\leq\beta\leq t,\ \int_\alpha^\beta F(t',t,x)\,dt'$$

exists as a Lebesgue integral, for all values of x, $(m<x<l)$, and for all those values of t such that $t-\mu>\alpha,\beta$; and as x approaches $m+0$, the integral converges to zero, uniformly for all such values of t.

Let $F(t',t,x)$ be a function of $t-t'$, say $F(t-t',x)$ and satisfy the conditions;

(a) $\operatorname*{Limit}_{x=m-0}\int_0^\mu F(z,x)\,dz = 1$

(b) $\int_0^\mu |F(z,x)|\,dz < A$

where A is independent of μ and x, for all sufficiently small values of $\mu>0$.

(I) $\operatorname*{Limit}_{x=m+0} F(t-t',x)=0$, when $t'\neq t$.

(II) $zF(z,x)$ has a total variation in the interval $(0,\mu)$ less than a fixed number independent of x.

---

* Hobson, *Theory of Functions of a Real Variable*, Vol. 2 (1926), pp. 443, 447 and 452.

The conclusion of the theorem is, that if these six conditions are satisfied by $F(t-t', x)$,

$$\operatorname*{Limit}_{x=m+0} \int_0^t F(t-t', x)\phi(t')\,dt' = \phi(t)$$

whenever

$$\int_0^{t'} \{\phi(t+t'') - \phi(t)\}\, dt''$$

has a differential coefficient, with respect to t', equal to zero for $t' = 0$. This is the case for almost all values of t, since $\phi(t)$ is summable.

We can easily show that the function

$$b/2\pi^{1/2}(t-t')^{-3/2}(x-m)e^{-b^2(x-m)^2/4(t-t')}$$

satisfies these conditions and so

$$\operatorname*{Limit}_{x=m+0} b/2\pi^{1/2} \int_0^t (t-t')^{-3/2}(x-m)e^{-b^2(x-m)^2/4(t-t')}\phi(t')\,dt' = \phi(t)$$

almost everywhere, if $\phi(t)$ is summable.

Similarly we can show

$$\operatorname*{Limit}_{x=m-0} (-a/2\pi^{1/2}) \int_0^t (t-t')^{-3/2}(x-m)e^{-a^2(x-m)^2/4(t-t')}\phi(t')\,dt' = \phi(t)$$

almost everywhere, if $\phi(t)$ is summable.

Applying these results to equation (2. 12) we have

$$\operatorname*{Limit}_{x=m-0} \partial u_1(x,t)/\partial x = a\pi^{1/2}\psi_1(t) - \int_0^t a^2 m(t-t')^{-3/2}\, e^{-a^2/(t-t')}\psi_1(t')\,dt'$$

and similarly from (2. 13) we have

$$\operatorname*{imit}_{=m+0} \partial u_2(x,t)/\partial x = -b\pi^{1/2}\psi_2(t) - \int_0^t [\tfrac{1}{2}b^2(m-l)(t-t')^{-3/2}]\,e^{-\beta^2/(t-t')}\psi_3(t')\,dt'.$$

In virtue of equation (2. 9) we have then

$$\operatorname*{Limit}_{x=m-0} K_1(\partial u_1/\partial x)(x,t) = \operatorname*{Limit}_{x=m+0} K_2(\partial u_2/\partial x)(x,t)$$

and the boundary condition (2. 5) is satisfied.

We need now to prove that

$$\operatorname*{Limit}_{x=x_0\pm 0} \int_{t_1}^{t_2} |\,\partial u_i/\partial x)(x,t)\,|\,dt = \int_{t_1}^{t_2} \operatorname*{Limit}_{x=x_0\pm 0} |\,\partial u_i/\partial x(x,t)\,|\,dt,$$

x_0 being the x— coördinate of any boundary.

As a typical case we consider

$$\underset{x=-m+0}{\text{Limit}} \int_{t_1}^{t_2} |\,(\partial u_1/\partial x)\,(x,t)\,|\,dt$$

$$= \underset{x=-m+0}{\text{Limit}} \int_{t_1}^{t_2} \Big| -a^2/2 \int_0^t (t-t')^{-3/2}(x-m)\,e^{-a^2(x-m)^2/4(t-t')}\psi_1(t')\,dt'$$

$$-a^2/2 \int_0^t (t-t')^{-3/2}(x+m)\,e^{-a^2(x+m)^2/4(t-t')}\psi_1(t')\,dt' \Big|\,dt.$$

The term

$$G(x,t) = -a^2/2 \int_0^t (t-t')^{-3/2}(x-m)\,e^{-a^2(x-m)^2/4(t-t')}\psi_1(t')\,dt$$

is bounded and continuous in x as x approaches $-m$, so that

$$(\alpha) \qquad \underset{x=-m+0}{\text{Limit}} \int_{t_1}^{t_2} |\,G(x,t)\,|\,dt = \int_{t_1}^{t_2} \underset{x=-m+0}{\text{Limit}}\,|\,G(x,t)\,|\,dt.$$

Consider the term

$$H(x,t) = -a^2/2 \int_0^t (t-t')^{-3/2}(x+m)\,e^{-a^2(x+m)^2/4(t-t')}\psi_1(t')\,dt'.$$

We define

$$\bar{H}(x,t) = a^2/2 \int_0^t (t-t')^{-3/2}(x+m)\,e^{-a^2(x+m)^2/4(t-t')}\,|\,\psi_1(t')\,|\,dt'$$

and by the theorem of Hobson quoted above

$$\underset{x=-m+0}{\text{Limit}}\ \bar{H}(x,t) = (\pi^{1/2}/a)\,|\,\psi_1(t)\,|$$

nearly everywhere, since $|\,\psi_1(t)\,|$ is summable.

We have $\bar{H}(x,t) \geqq |\,H(x,t)\,|$ for all x and t.

Consider now

$$\underset{x=-m+0}{\text{Limit}} \int_0^T \bar{H}(x,t)\,dt$$

$$= \underset{x=-m+0}{\text{Limit}} \int_0^T dt\,(a^2/2) \int_0^t (t-t')^{-3/2}(x+m)\,e^{-a^2(x+m)^2/4(t-t')}\,|\,\psi_1(t')\,|\,dt'.$$

By a change of order of integration we have

$$\int_0^T dt\,(a^2/2) \int_0^t (t-t')^{-3/2}(x+m)\,e^{-a^2(x+m)^2/4(t-t')}\,|\,\psi_1(t')\,|\,dt'$$

$$= \int_0^T dt'\,|\,\psi_1(t')\,|\,a^2/2 \int_{t'}^T (t-t')^{-3/2}(x+m)\,e^{-a^2(x+m)^2/4(t-t')}\,dt$$

$$= \int_0^T g(t',T,x)\,|\,\psi_1(t')\,|\,dt'$$

where *

$$g(t', T, x) = a^2/2 \int_{t'}^{T} (t - t')^{-3/2} (x + m) e^{-a^2(x+m)^2/4(t-t')} dt, \qquad x > -m$$

$$\leq a^2/2 \int_{t'}^{\infty} (t - t')^{-3/2} (x + m)\, e^{-a^2(x+m)^2/4(t-t')} dt$$

Putting $\xi = a(x + m)/2(t - t')^{1/2}$ we have

$$g(t', T, x) \leq 2a \int_0^{\infty} e^{-\xi^2} d\xi = a\pi^{1/2} \text{ independent of } t', T, x.$$

We have then

$$\operatorname*{Limit}_{x=-m+0} \int_0^T g(t', T, x) \mid \psi_1(t') \mid dt = \int_0^T \operatorname*{Limit}_{x=-m+0} g(t', T, x) \mid \psi_1(t') \mid dt'$$

$$= a\pi^{1/2} \int_0^T \mid \psi_1(t') \mid dt = \int_0^T \operatorname*{Limit}_{x=-m+0} \bar{H}(x, t')\, dt'$$

That is

$$\operatorname*{Limit}_{x=-m+0} \int_0^T \bar{H}(x, t)\, dt = \int_0^T \operatorname*{Limit}_{x=-m+0} \bar{H}(x, t)\, dt$$

And, since $\bar{H}(x, t) \geq \mid H(x, t) \mid \geq 0$, and since Limit $H(x, t)$ exists nearly everywhere, we have

$$\operatorname*{Limit}_{x=-m+0} \int_0^T \mid H(x, t) \mid dt = \int_0^T \operatorname*{Limit}_{x=-m+0} \mid H(x, t) \mid dt$$

Hence recalling (α) we have

$$\operatorname*{Limit}_{x=-m+0} \int_0^T \mid (\partial u_1/\partial x)(x, t) \mid dt = \int_0^T \operatorname*{Limit}_{x=-m+0} \mid (\partial u/\partial x)(x, t) \mid dt$$

as was to be shown.†

Thus the solution given by $u_1(x, t)$ and $u_2(x, t)$ in the inner and outer regions respectively satisfies the conditions of the Uniqueness Theorem B.

It remains to be shown that $\psi_1(t)$, $\psi_2(t)$ and $\psi_3(t)$ can be found to satisfy equations (2.15) to (2.17). This will be done in the next two sections.

---

* For $T > t'$ $\operatorname*{Limit}_{x=-m+0} g(t', T, x) = a\pi^{1/2}$.

† We have used the limits 0 and T, the equation holds for any limits t_1 and t_2.

PART III. Case of Infinite Side Regions.

A case of considerable interest is the one where the outer regions are infinite.* In this section we give a method of solution of this case.

Here we have no outer boundary and replace the condition (2.3) by the condition that $\lim_{x=\infty} [\partial u_2(x, t)/\partial x] = 0$. The solution for the outer region differs from that above by the absence of the term involving $\psi_3(t)$. The result is that we have two integral equations for the two functions $\psi_1(t)$ and $\psi_2(t)$. These are

$$(3.1) \qquad \psi_2(t) = -c\psi_1(t) + ca\pi^{-1/2} \int_0^t (t-t')^{-3/2} e^{-a^2/(t-t')} \psi_1(t')\, dt'$$

and

$$(3.2) \qquad \int_0^t \{(t-t')^{-\frac{1}{2}} e^{-a^2/(t-t')} + (t-t')^{-1/2}\} \psi_1(t')\, dt' = f_1 + \int_0^t (t-t')^{-1/2} \psi_2(t')\, dt'$$

If we substitute the value of $\psi_2(t)$ given by the right-hand member of (3.1) into (3.2), change the order of integration and carry out the integration in the inner integrals, we have

$$(3.3) \qquad \int_0^t \{(t-t')^{-1/2}[1 + (\overline{1-c}/\overline{1+c}) e^{-a^2/(t-t')}]\} \psi_1(t')\, dt' = f_1/(1+c)$$

This is an equation of the sort dealt with in Part I with

$$K(t, t') = (\overline{1-c}/\overline{1+c})\, e^{-a^2/(t-t')}$$

where c is a positive constant. Since $K(t, t')$ vanishes exponentially for $t = t'$ and is bounded elsewhere, the conditions of Part I are satisfied and we can find a summable solution of (3.3) by the method given in that part. This solution, $\psi_1(t)$, substituted in (3.1) gives a summable function $\psi_2(t)$. These substituted in (2.9) and the modified (2.7) give functions $u_1(x, t)$ and $u_2(x, t)$ that solve the problem.

PART IV. Solution of Equations for Finite Side Regions.

In this section we consider the more general case of finite outer regions. As in Part III, we do not solve for $\psi_1(t)$, $\psi_2(t)$ and $\psi_3(t)$ directly, but, assuming the existence of summable solutions, $\psi_1(t)$, $\psi_2(t)$ and $\psi_3(t)$, for

---

* This problem is treated by A. Sommerfeld, "Zur analytischen Theorie der Wärmeleitung," *Mathematische Annalen*, Vol. 45 (1894), p. 270, by constructing the Green's function by the method of images.

$\int_0^{t''}\psi_1(t)\,dt$, $\int_0^{t''}\psi_2(t)\,dt$ and $\int_0^{t''}\psi_3(t)\,dt$. In order to do this we must make changes in the equations (2.15), (2.16) and (2.17) so that they will involve the integrals.

We multiply equation (2.15) by $(t''-t)^{-1/2}$ and integrate from $t=0$ to $t=t''$, thus

$$(4.1)\quad \int_0^{t''}(t''-t)^{-1/2}dt\int_0^{t}(t-t')^{-1/2}\psi_3(t')\,dt' = \int_0^{t''}(t''-t)^{-1/2}f_2(t)\,dt$$
$$-\int_0^{t''}(t''-t)^{-1/2}dt\int_0^{t}(t-t')^{-1/2}e^{-\beta^2/(t-t')}\psi_2(t')\,dt'$$

The first term by a change of order of integration gives

$$\pi\int_0^{t''}\psi_3(t')\,dt'.$$

Integrating the inner integral of the last term by parts gives

$$[(t-t')^{-1/2}e^{-\beta^2/(t-t')}\int_0^{t'}\psi_2(t'')\,dt'']_{t'=0}^{t'=t}$$
$$-\int_0^{t}\left\{\int_0^{t'}\psi_2(t'')\,dt''\right\}[(t-t'-2\beta^2)/2\,(t-t')^{5/2}]e^{-\beta^2/(t-t')}dt'$$

The term outside the sign of integration is zero. Putting this in and changing the order of integration gives, in place of (4.1)

$$(4.2)\quad \pi\int_0^{t''}\psi_3(t')\,dt' = \int_0^{t''}(t''-t')^{-1/2}f_2(t')\,dt'$$
$$+\int_0^{t''}\left\{\int_0^{t'}\psi_2(t)\,dt\right\}dt'\int_{t'}^{t''}[(t-t'-2\beta^2)/2(t-t')^{5/2}]e^{-\beta^2/(t-t')}dt'$$
$$=\int_0^{t''}(t''-t')^{-1/2}f_2(t')\,dt'-\int_0^{t''}\left\{\int_0^{t'}\psi_2(t)\,dt\right\}[\beta/(t''-t')^{3/2}]\;e^{-\beta^2/(t''-}$$

We integrate the equation (2.16) from $t=0$ to $t=t''$ and in the last two terms integrate by parts and then change the order of integration, this gives

$$(4.3)\quad \int_0^{t''}\psi_2(t')\,dt' = -c\int_0^{t''}\psi_1(t')\,dt'$$
$$-c\alpha\pi^{-1/2}\int_0^{t''}\left\{\int_0^{t'}\psi_3(t)\,dt\right\}dt'\int_0^{t''}(\partial/\partial t')[(t-t')^{-3/2}e^{-\alpha^2/(t''-t')}]\,dt$$
$$-\beta\pi^{-1/2}\int_0^{t''}\left\{\int_0^{t'}\psi_1(t)\,dt\right\}dt'\int_{t'}^{t''}(\partial/\partial t')[(t-t')^{-3/2}e^{-\beta^2/(t''-t')}]\,dt$$

$$= - c \int_0^{t''} \psi_1(t')\,dt'$$

$$+ c\alpha\pi^{-1/2} \left\{ \int_0^{t''} \int_0^{t'} \psi_1(t)\,dt \right\} (t'' - t)^{-3/2} e^{-\alpha^2/(t''-t')}\,dt'$$

$$+ \beta\pi^{-1/2} \int_0^{t''} \left\{ \int_0^{t''} \psi_3(t)\,dt \right\} (t'' - t')^{-3/2} e^{-\beta^2/(t''-t')}\,dt.$$

Multiplying equation (2.17) by $(t'' - t)^{-1/2}$ and integrating from $t = 0$ to $t = t''$, we obtain, in a manner similar to that used in deriving equation (4.2),

$$\alpha \int_0^{t''} \left\{ \int_0^{t'} \psi_1(t)\,dt \right\} (t'' - t')^{-3/2} e^{-\alpha^2/(t''-t')}\,dt' \tag{4.4}$$

$$+ \pi \int_0^{t''} \psi_1(t')\,dt' = 2f_1(t'')^{-1/2} + \pi \int_0^{t''} \psi_2(t')\,dt'$$

$$+ \beta \int_0^{t''} \left\{ \int_0^{t'} \psi_3(t)\,dt \right\} (t'' - t')^{-3/2} e^{-\beta^2/(t''-t')}\,dt'.$$

By putting

$$F(t'') = \int_0^{t''} (t'' - t')^{-1/2} f_2(t')\,dt'$$

$$V_i(t'') = \int_0^{t''} \psi_i(t)\,dt \qquad (i = 1, 2, 3)$$

$$H_\rho(t'', t') = [\rho/(t'' - t')^{3/2}]\, e^{-\rho^2/(t''-t')} \qquad (\rho = \alpha, \beta)$$

we can write (4.2), (4.3) and (4.4) as

$$\pi V_3(t'') = F(t'') - \int_0^{t''} V_2(t') H_\beta(t'', t')\,dt' \tag{4.5}$$

$$V_2(t'') = - c\, V_1(t'') + h \int_0^{t''} V_1(t') H_\alpha(t'', t')\,dt' \tag{4.6}$$

$$+ \pi^{-1/2} \int_0^{t''} V_3(t') H_\beta(t'', t')\,dt'$$

and

$$\int_0^{t''} V_1(t') H_\alpha(t'', t')\,dt' + \pi V_1(t'') \tag{4.7}$$

$$= 2f_1(t'')^{-1/2} + \pi V_2(t'') + \int_0^{t''} V_3(t') H_\beta(t'', t')\,dt'.$$

Equation (4.6) multiplied by π and added to equation (4.7) gives

$$V_1(t'') = 1/\pi(1 + c) \{2f_1(t'')^{1/2} \tag{4.8}$$

$$+ (c\pi^{1/2} - 1) \int_0^{t''} V_1(t') H_\alpha(t'', t')\,dt'$$

$$+ (1 + \pi^{1/2}) \int_0^{t''} V_3(t') H_\beta(t'', t')\,dt'\}.$$

Equation (4.6) multiplied by π and subtracted from c times equation (4.7) gives

$$(4.9)\qquad V_2(t'') = 1/\pi(1+c)\{-2cf_1(t'')^{1/2} + c(1+\pi^{1/2})\int_0^{t''} V_1(t')H_\alpha(t'',t')dt' - (c-\pi^{1/2})\int_0^{t''} V_3(t')H_\beta(t'',t')dt'\}$$

We have in equations (4.5), (4.8) and (4.9) a system of Volterra integral equations of the second kind with bounded kernels. Such a system is solvable by a process of successive approximations that *necessarily* converges.* It may be pointed out that, since the kernels involved vanish exponentially, the process of approximation converges rapidly.

The terms on the right-hand sides of equations (4.5), (4.8) and (4.9) are of the types shown to be absolutely continuous in Part I, and so the functions $V_1(t)$, $V_2(t)$ and $V_3(t)$ are absolutely continuous. They thus possess derivatives nearly everywhere and are the integrals of these derivatives. We wish to show that the functions $\psi_1(t) = V_1'(t)$, $\psi_2(t) = V_2'(t)$ and $\psi_3(t) = V_3'(t)$ satisfy the equations (2.15), (2.16) and (2.17).

We do this by a device similar to that employed in Part I. That is we substitute $V_2'(t)$ for $\psi_2(t)$ and $V_3'(t)$ for $\psi_3(t)$ in the right- and left-hand members of equation (2.15) and call the difference between the two members $D_1(t)$. Then multiplying by $(t''-t)^{-1/2}$ and integrating from $t=0$ to $t=t''$, changing order of integration and integrating by parts as was done to obtain equations (4.1) and (4.2) we have

$$\int_0^{t''}(t''-t)^{-1/2}D(t)\,dt = \pi\int_0^{t''}V_3'(t)\,dt - F(t'') + \beta\int_0^{t''}\left\{\int_0^{t''}V_2'(t)\,dt\right\}H_\beta(t'',t')\,dt'$$

which is zero since $\int_0^{t'}V_2'(t)\,dt = V_2(t')$ and $\int_0^{t'}V_3'(t)\,dt = V_3(t')$ and satisfy equation (4.6). Hence $D_1(t)$ is zero almost everywhere and equation (2.15) is satisfied.

Similarly equations (2.16) and (2.17), by the methods used to obtain equations (4.3) and (4.5), are shown to be satisfied almost everywhere in virtue of the equations (4.9) and (4.10).

---

* See Volterra, *Leçons sur les Équations Intégrales*, p. 71.

PART V. EXTENSION TO GENERAL CASES.

The method of Part IV will apply to any number of regions and to conditions which are not symmetric. The only merit of the symmetric conditions is that they reduce the number of conditions by half.

The initial temperatures were taken to be constant in order to make the form of equations (2. 7) and (2. 9) as simple as possible. If the initial temperatures are not constant, the constant u_1 in equation (2. 9) is replaced by the expression

$$a/2\pi^{1/2} \int_{-m}^{+m} u_1(x')\, t^{-1/2}\, e^{-a^2(x-x')^2/4t}\, dx'$$

where $u_1(x)$ is the initial temperature. This expression satisfies equation (2. 1) and approaches $u_1(x)$, as t approaches zero, nearly everywhere if $u_1(x)$ is summable.* For varying initial temperature in the outer regions the constant u_2 is replaced by an analogous expression involving $u_2(x)$, the initial temperature in the outer region. We can show that this solution satisfies the conditions of the Uniqueness Theorem B if $u_1(x)$ and $u_2(x)$ are bounded. The introduction of these expressions necessitates the use of equations more complicated than (2. 8) to (2. 10), but makes no essential difference in the method.

In closing I should like to acknowledge my indebtedness to Dr. G. C. Evans, both for his proposal of this problem and for his timely criticisms and suggestions.

THE RICE INSTITUTE,
HOUSTON, TEXAS.

* Hobson, *loc. cit.*, Vol. 2, p. 402.

Acc. No. ....
Book No. ....
LIBRARY.

COMPLEMENTS OF POTENTIAL THEORY.*

By Griffith C. Evans.

Introduction. Let C be a simple, closed, rectifiable plane curve, and T_+, T_- the interior and exterior regions, respectively, bounded by the curve. Let Q be a fixed and P a variable point on C, and let $\nu(e)$ be an additive function of point sets e on C, defined for all sets e, measurable in the Borel sense, with respect to arc length s of C. We are to consider the limiting values of potentials of a double layer, and the derivatives of potentials of a single layer, formed for these general mass distributions $\nu(e)$, as the point M at which they are given approaches Q from T_+ or from T_-.

In order to obtain these limiting values a restriction on the nature of the curve in the neighborhood of the particular point Q is unavoidable (see § 2, below). It is essentially a smoothing of the curve which is necessary, rather than a smoothing of the mass distribution, as the examples which are to be given will show.

The coördinates x, y of P are absolutely continuous functions of s, and the quantities $\cos(x, s)$, $\cos(y, s)$ are defined almost everywhere on C, and are bounded, in numerical value ≤ 1. Moreover, irrespective of their definition on the remaining set of zero measure, they are summable, and the formulae †

$$x_{P_2} - x_{P_1} = \int_{P_1}^{P_2} \cos(x, s)\, ds, \qquad y_{P_2} - y_{P_1} = \int_{P_1}^{P_2} \cos(y, s)\, ds$$

are valid, for any two points P_1, P_2 on C.

Lemma I. *Let $P_1, P_2, \cdots$ be an infinite sequence of points, lying ahead of P on C, with $\lim P_k = P$, and suppose that the direction of the ray PP_k has a corresponding limit which makes an angle θ with a fixed direction. If for some neighborhood of P (exclusive of P), measured from P in the positive sense on C, the direction θ' of the curve, wherever it exists, satisfies*

* Presented in part to the American Mathematical Society, September, 1931, under the title: "The Cauchy Integral and Summable Boundary Values." The part of Theorem I, below, relating to $u(M)$, $u_1(M)$ has been proved for surfaces in Evans and Miles, "Potentials of General Masses in Single and Double Layers," *American Journal of Mathematics*, Vol. 53 (1931), pp. 493-516, but only in the case where approach is along the normal to the surface.

† Lebesgue, *Leçons sur l'intégration*, Paris (1928), pp. 198-201.

the inequality $\theta_1 < \theta' < \theta_2$, *with* $\theta_2 - \theta_1 \leq \pi$ *then* θ *satisfies the inequality* $\theta_1 \leq \theta \leq \theta_2$.

We take $\theta_1 = 0$, and that direction as the axis of x; and we take the origin at P. We assume first that $\theta_2 \leq \pi/2$. We have, almost everywhere,

$$0 < \cos(y, s) < \sin \theta_2 \leq 1, \quad 0 \leq \cos \theta_2 < \cos(x, s) < 1, \quad \text{for} \quad 0 < s \leq \delta,$$

and from the integral formulae given above,

$$0 < y_k < s_k \sin \theta_2, \qquad x_k > s_k \cos \theta_2 \geq 0, \qquad 0 < y_k/x_k < \tan \theta_2,$$

from which the truth of the Lemma is evident.

Suppose finally that $\pi/2 < \theta_2 \leq \pi$. We still have $y_k > 0$, and therefore $\theta \geq 0$. If $\theta \leq \pi/2$ the proposition is verified. We suppose then that from a certain k we have $\theta_k > \pi/2$. Let s'_k denote the measure of the portion of the arc PP_k for which $\cos(x, s)$, where it exists, is negative. Then

$$y_k > s'_k \sin \theta_2 \geq 0, \qquad |\, x_k \,| < s'_k \,|\cos \theta_2\,|, \qquad y_k/\,|\, x_k \,| > |\tan \theta_2\,| \geq 0,$$

from which the truth of the Lemma is again evident. We cannot however extend the proposition to the case where $\theta_2 - \theta_1 > \pi$. A similar statement holds of course for approach to P along C in the positive sense.

If we are dealing with, say, the potential of a double layer due to an absolutely continuous mass function—that is, a density distribution—on C, it is unnecessary to say anything about the points where a direction of the curve fails to exist; for these constitute a set of measure zero on C, and the integral which gives the potential is not affected by changing the doublets on this set. On the other hand, if we have a general distribution of mass as a double layer, it may even happen that a doublet of positive mass will be placed at some single point of the curve where there is no tangent direction, and the value of the potential will involve a not-zero term depending on the orientation of the doublet at this point. We must therefore, in specifying a distribution of a double layer, define an orientation of the doublet at every point of C—at least, at every point where the density is infinite—whether or not the curve itself has a direction at that point. We do this by assigning an approximate direction to the curve.

Let P be an arbitrary point of C, and P' a point nearby at a distance ϵ from P in the positive sense around C; let θ_ϵ be the angle made by PP' with a fixed x-direction. For a given ϵ, θ_ϵ is a continuous function of P, and the superior and inferior limits

$$\overline{\lim_{\epsilon=0}}\ \theta_\epsilon = \bar{\theta}(P), \qquad \underline{\lim_{\epsilon=0}}\ \theta_\epsilon = \underline{\theta}(P)$$

are functions measurable in the Borel sense on C.

The set of points where one of the quantities $\bar{\theta}(P)$, $\underline{\theta}(P)$ is infinite, is measurable in the Borel sense, and therefore the following function is likewise measurable in that sense:

$$\theta_+(P) = (1/2)[\bar{\theta}(P) + \underline{\theta}(P)] - 2n(P)\pi, \quad \text{if } \bar{\theta}(P),\ \underline{\theta}(P) \text{ are finite,}$$
$$n(P) \text{ an integer, chosen so that } 0 \leq \theta_+(P) < 2\pi,$$
$$= 0, \quad \text{if } \bar{\theta}(P) \text{ or } \underline{\theta}(P) \text{ is not finite.}$$

Similarly a limiting direction $\theta_-(P)$ of $P'P$ is chosen, for approach of P' to P in a positive sense along C. The function

$$\theta(P) = (1/2)[\theta_+(P) + \theta_-(P)]$$

is then bounded, measurable in the Borel sense, and $\tan\theta = dy/dx$, $\cos\theta = dx/ds$, $\sin\theta = dy/ds$, wherever these quantities exist and satisfy the condition $(dx/ds)^2 + (dy/ds)^2 = 1$.

For convenience, we call the direction indicated by $\theta(P)$ the direction of C at P, and fix the direction of the doublet at P in the customary way with reference to this direction. Definitions which make use of averages would of course also be possible, but they are not necessary here.

Let M be a point of T_+ or T_-, $r = MP$ (with direction MP when angles are to be specified), n_P the direction of the normal at P, taken towards T_+, and perpendicular to the direction, as just defined, of the curve at P. Then the quantities

$$\cos\phi = \cos(r, n_P) = \cos(x, n_P)\cos(x, r) + \cos(y, n_P)\cos(y, r)$$
$$\sin\phi = \sin(r, n_P) = \cos(r, s_P)$$

are everywhere defined on C, measurable in the Borel sense, and in absolute value ≤ 1. The quantities

$$\cos\phi' = \cos(r, n_Q), \qquad \sin\phi' = \sin(r, n_Q)$$

with a fixed Q, are defined for all P on C, are continuous, and in numerical value ≤ 1.

The four Stieltjes integrals

$$\text{(A)} \qquad u(M) = \int_C r^{-1}\cos\phi\, d\nu(e_P)$$

$$\text{(B)} \qquad u_1(M) = \int_C r^{-1}\cos\phi'\, d\nu(e_P)$$

$$\text{(C)} \qquad u_2(M) = \int_C r^{-1}\sin\phi\, d\nu(e_P)$$

$$\text{(D)} \qquad u_3(M) = \int_C r^{-1}\sin\phi'\, d\nu(e_P)$$

are now all defined for M in T_+ or in T_-. The integrals (B) and (D) are given in the usual sense; those in (A) and (C) are Radon * or Daniell † generalized integrals. The (A) and (C) may be regarded as potentials due to general mass distributions of doublets, respectively normal and tangent to the curve C; the integrals (B) and (D) are the derivatives in the fixed direction n_Q and perpendicular to it, respectively, of a general simple mass distribution on C.

A general mass distribution of doublets along C, whose orientation is determined at points of C by direction cosines which are measurable in the Borel sense with respect to s, may be replaced, as far as its potential is concerned, by two integrals of forms (A) and (C), with different mass functions $\nu_1(e)$, $\nu_2(e)$ respectively. For if $h(P)$ is bounded and measurable (B) the integral

$$\int_Q^P h(P')\,d\nu(e_{P'})$$

is a function of bounded variation on C and determines such a distribution of mass.

In conclusion to these introductory considerations, we introduce the following condition (γ) with respect to the point Q. We make the convention that the symbol s merely denotes numerical value, except when, in conjunction with some other letter, it forms part of the symbol for an angle.

(γ). *Given the fixed point Q and the variable point P, on C, there subsists the inequality $|\sphericalangle(n_Q, n_P)| < \gamma(s)$, s the arc QP, where $\gamma(s)$, for $s > 0$, is a positive, monotonic (non-decreasing) function, such that the integral*

$$\int_0^s \gamma(s)/s\,ds = m(s)$$

is convergent.

We have $\lim m(s) = 0$ if $\lim s = 0$.

It is not necessary to assume that $\gamma(s)$ is continuous. Moreover, if the condition (γ) is assumed to hold merely for points P in the neighborhood of Q where dx/ds and dy/ds are defined in the usual way, it will hold throughout, with the extended definition of the direction of C, provided that we write $\gamma(s) = \gamma(s+0)$. For that is the statement expressed in Lemma I.

* J. Radon, "Theorie und Anwendung der absolut additiven Mengenfunktionen," *Sitzungsberichte der Akademie der Wissenschaften*, Wien (1913), p. 1295.

† P. J. Daniell, (I) "A General Form of Integral," *Annals of Mathematics*, Vol. 19 (1918), pp. 279-294; (II) "Further Properties of the General Integral," *ibid.*, Vol. 21 (1920), pp. 203-220.

The following theorem will be proved.

THEOREM I. *Let the condition* (γ) *be satisfied at the point* Q *of* C, *and let* $\nu(e)$ *have at* Q *a unique derivative** $\nu'(e) = A$, *with respect to* s. *Denote by* M' *the image of* M *by reflexion across the tangent line at* Q. *Then as* M *in* T_+ *approaches* Q ($\lim\limits_{M=Q+}$) *or as* M *in* T_- *approaches* Q ($\lim\limits_{M=Q-}$), *in the wide sense,*† *the following relations hold*

$$\text{(E)} \qquad \lim_{M=Q\pm} u(M) = \mp 2\pi A + \int_C \frac{\cos(QP, n_P)}{QP}\, d\nu(e_P)$$

$$\text{(E}_1\text{)} \qquad \lim_{M=Q\pm} u_1(M) = \mp 2\pi A + \int_C \frac{\cos(QP, n_Q)}{QP}\, d\nu(e_P)$$

$$\text{(E}_2\text{)} \qquad \lim_{M=Q\pm} \{u_2(M) - u_2(M')\} = 0$$

$$\text{(E}_3\text{)} \qquad \lim_{M=Q\pm} \{u_3(M) - u_3(M')\} = 0.$$

1. *Lemmas on Stieltjes integrals.* Although special cases of the following proposition are frequently used, it is convenient to have the statement in explicit form. The integrals are generalized Stieltjes integrals.

LEMMA II. *Let* $f(x)$ *be bounded and measurable in the Borel sense, and* $g_1(x)$, $g_2(x)$ *of bounded variation,* $a \leq x \leq b$; *then* $g(x) = g_1(x)g_2(x)$ *is of bounded variation, and*

$$(1.1) \qquad \int_a^b f(x)\,dg(x) = \int_a^b f(x)g_1(x)\,dg_2(x) + \int_a^b f(x)g_2(x)\,dg_1(x).$$

The boundedness of the variation of $g(x)$ is of course well known. It is a consequence of the identity

$$g(x_{i+1}) - g(x_i) = g_1(x_{i+1})\{g_2(x_{i+1}) - g_2(x_i)\} + g_2(x_i)\{g_1(x_{i+1}) - g_1(x_i)\}.$$

When $f(x)$, $g_1(x)$, $g_2(x)$ are, in addition, all continuous, (1.1) follows immediately from the definition of the Riemann-Stieltjes integral. In fact

$$\Sigma f(x_i)\{g(x_{i+1}) - g(x_i)\} = \Sigma f(x_i)g_1(x_{i+1})\{g_2(x_{i+1}) - g_2(x_i)\} + \Sigma f(x_i)g_2(x_i)\{g_1(x_{i+1}) - g_1(x_i)\}.$$

---

* That is, a unique derivative with respect to regular families.

† Let Q be a point of C and Δ_1 a circular sector with vertex at Q, but otherwise contained, with its boundary, in T_+ (T_-). Let Δ_2 be a circular sector with vertex at Q, but otherwise contained, with its boundary, in the interior of Δ_1. We say that M approaches Q in the wide sense from T_+ (T_-) if, approaching Q, it comes and remains eventually within some Δ_2.

The lemma is then established by passages to the limit with (1.1) for continuous functions as a basis.

For this purpose we write $g_{nm}(x) = g_{1n}(x) g_{2m}(x)$ where $g_{ik}(x)$, $(i = 1, 2)$, is constructed so as to have the following properties:

(a) $g_{ik}(x)$ is continuous, $a \leq x \leq b$.

(b) $\lim\limits_{k=\infty} g_{ik}(x) = g_i(x)$, $a \leq x \leq b$.

(c) $| g_{ik}(x_2) - g_{ik}(x_1) | \leq t(x_2) - t(x_1)$, $a \leq x_1 < x_2 \leq b$, where $t(x)$ is of bounded variation.

The property (c) implies the following:

(d) $g_{ik}(x)$ is of bounded variation uniformly for all k.

This construction can be made by a polygonal approximation.

Consider for instance a function of bounded variation $\alpha(x) = p(x) - n(x) + \alpha(a)$ where $p(x)$ and $-n(x)$ are respectively positive and negative variation functions for $\alpha(x)$. For $\alpha_m(x)$ we take $p_m(x) - n_m(x)$, where $p_m(x)$ is a polygonal function, formed by joining successively with straight lines a finite number of points on the graph of $p(x)$, the vertices for $p_{m+1}(x)$ to be obtained by adding new ones to those of $p_m(x)$ in such a way that as m tends to ∞ they eventually become dense in (a, b) and include the points for which $x = a$ and $x = b$, and all the points of discontinuity of $p(x)$. A similar construction is used for $n_m(x)$. Evidently then lim $(m = \infty)$ $\alpha_m(x) = \alpha(x)$, $\alpha_m(x)$ is continuous, and

$$| \alpha_m(x_2) - \alpha_m(x_1) | \leq t(x_2) - t(x_1)$$

where $t(x)$ is the total variation function of $\alpha(x)$, namely $p(x) + n(x)$.

Let $f(x)$ remain continuous. Equation (1.1) holds when $g_1(x)$, $g_2(x)$ and $g(x)$ are replaced respectively by $g_{1n}(x)$, $g_{2m}(x)$ and $g_{nm}(x)$. From (a), (b), (d) it follows, as a well known theorem,* that

$$\lim_{n=\infty} \int_a^b f(x) g_{2m}(x) \, dg_{1n}(x) = \int_a^b f(x) g_{2m}(x) \, dg_1(x)$$

$$\lim_{n=\infty} \int_a^b f(x) \, dg_{nm}(x) = \int_a^b f(x) \, dg_m(x)$$

where $g_m(x) = \lim$ $(n = \infty)$ $g_{nm}(x)$ is of bounded variation, uniformly in m, and lim $(m = \infty)$ $g_m(x) = g(x)$. Also

* Helly, *Sitzungsberichte der Akademie der Wissenschaften*, Wien (1912), p. 288; Radon, *ibid., loc. cit.*; Bray, *Annals of Mathematics*, Vol. 20 (1919), p. 180.

$$\lim_{n=\infty}\int_a^b f(x)g_{1n}(x)\,dg_{2m}(x)=\int_a^b f(x)g_1(x)\,dg_{2m}(x),$$

as a fundamental property of the general integral. Accordingly (1.1) holds when $f(x)$ is continuous, and $g_2(x)$ is replaced by $g_{2m}(x)$.

Let now m tend to ∞. We have, as above,

$$\lim_{m=\infty}\int_a^b f(x)\,dg_m(x)=\int_a^b f(x)\,dg(x)$$

$$\lim_{m=\infty}\int_a^b f(x)g_{2m}(x)\,dg_1(x)=\int_a^b f(x)g_2(x)\,dg_1(x),$$

and an extension by Daniell * of the theorem, referred to above, yields the fact that

$$\lim_{m=\infty}\int_a^b f(x)g_1(x)\,dg_{2m}(x)=\int_a^b f(x)g_1(x)\,dg_2(x).$$

Consequently (1.1) holds, under the hypotheses of the lemma, provided that $f(x)$ is continuous. That it is also valid when $f(x)$ is merely bounded and measurable in the Borel sense is an immediate consequence of the postulates which define the general integral.†

Given a function $\nu(x)$, of bounded variation, we say that it has a unique derivative, on regular families, at $x=0$, if the corresponding function of plurisegments has a unique derivative, on regular families, at $x=0$, or, what amounts to the same thing, if it is continuous at $x=0$, and if the corresponding additive function of point sets has such a derivative. A plurisegment r is the sum of a finite number or of a denumerable infinity of segments s, $s=(x',x'')$, no two of the segments having interior points in common; the measure of a plurisegment is the sum of the measures of its segments. We say that $\nu(r)$ corresponds to $\nu(x)$ if when s is the interval (x',x''), $\nu(s)=\nu(x'')-\nu(x')$. The theory of such functions is given by Vitali ‡ and Maria.§

* Daniell, *loc. cit.* (II), see 3(5a), p. 218. If $\lim\,(m=\infty)\ \alpha_m(x)=\alpha(x)$, $a\leq x\leq b$ and $|\alpha_m(x_2)-\alpha_m(x_1)|\leq\beta(x_2)-\beta(x_1)$ for all x_1, x_2, $a\leq x_1<x_2\leq b$, independently of m where $\beta(x)$ is of bounded variation, then

$$\lim_{m=\infty}\int_a^b f(x)\,d\alpha_m(x)=\int_a^b f(x)\,d\alpha(x),$$

provided that $f(x)$ is summable with respect to $\beta(x)$.

† Daniell, *loc. cit.* (I).

‡ G. Vitali, "Analisi delle funzioni a variazione limitata," *Rendiconti del Circolo Matematico di Palermo*, Vol. 46 (1922), p. 388.

§ A. Maria, "Functions of Plurisegments," *Transactions of the American Mathematical Society*, Vol. 28 (1926), pp. 448-471.

A sequence of plurisegments r_i forms a regular family about $x = 0$, with parameter of regularity k, if

$$\varliminf_{i=\infty} \frac{\text{meas. } r_i}{\omega_i} = k > 0.$$

where ω_i is the measure of the smallest segment with center $x = 0$ which contains r_i. The function $\nu(r)$ has a unique derivative A, on regular families, at $x = 0$, if

$$\lim_{i=\infty} \frac{f(r_i)}{\text{meas. } r_i} = A$$

for every regular family $\{r_i\}$ about $x = 0$.

The following lemma * will accordingly be proved if we prove the corresponding theorem for additive functions of plurisegments. It is also true when stated in terms of additive functions of point sets.

LEMMA III. *If $\nu(x)$ has the unique derivative A (on regular families) at $x = 0$, the total variation function $t(x)$ of $\nu(x)$ has the unique derivative $| A |$ at $x = 0$; moreover the positive variation function $p(x)$ has the unique derivative $\{ | A | + A\}/2$, and the negative variation function $-n(x)$ has the unique derivative $-\{ | A | - A\}/2$*, at $x = 0$.

Let r_i constitute a regular family of plurisegments about $x = 0$, with parameter of regularity k. We consider first $A > 0$, and show that

$$\lim (i = \infty)\ p(r_i)/\text{meas. } r_i = A.$$

Given r_i we can find a finite plurisegment f_i, composed of a finite number of the segments of r_i, which has the following properties:

$$\text{meas. } (r_i - f_i) < \delta_i, \quad \delta_i \text{ arbitrary}, > 0,$$
$$p(r_i) - p(f_i) < \eta_i, \quad \eta_i \text{ arbitrary}, > 0,$$
$$\varliminf_{i=\infty} \frac{\text{meas. } f_i}{\omega_i} = k,$$
$$\lim_{i=\infty} \left| \frac{p(r_i)}{\text{meas. } r_i} - \frac{p(f_i)}{\text{meas. } f_i} \right| = 0.$$

Hence it will be sufficient to show that $\lim p(f_i)/\text{meas. } f_i = A$.

Given f_i we can find a finite plurisegment f'_i, $f'_i \subset f_i$, such that

$$\nu(f'_i) \leq p(f_i) < \nu(f'_i) + \epsilon_i, \qquad (\epsilon_i > 0), \tag{1.2}$$

* Proved also by my colleague H. E. Bray, some time ago, but not published. It applies of course in n dimensions.

and if we let $f_i'' = f_i - f'_i$, f_i'' will be a finite plurisegment, and

$$\nu(f'_i) + \nu(f_i'') = p(f_i) - n(f_i) = \nu(f_i).$$

Hence, by subtraction,

$$(1.3) \qquad -\nu(f_i'') \leq n(f_i) < -\nu(f_i'') + \epsilon_i.$$

We take ϵ_i small, so that $\lim\ (i = \infty)\ \epsilon_i/\omega_i = 0$.

We note first that the family f_i'' can contain no subsequence which constitutes a regular family with respect to the intervals ω_i, that is, allow the selection of a subsequence f_j'' such that

$$\lim_{j=\infty} \frac{\text{meas. } f_j''}{\omega_j} = k'' > 0.$$

For in this case,

$$\frac{n(f_j)}{\text{meas. } f_j} \leq -\frac{\nu(f_j'')}{\text{meas. } f_j''} \frac{\text{meas. } f_j''}{\omega_j} \frac{\omega_j}{\text{meas. } f_j} + \frac{\epsilon_j}{\omega_j} \frac{\omega_j}{\text{meas. } f_j}.$$

As j tends to ∞, $\nu(f_j'')/\text{meas. } f_j''$ approaches A, and the first term in the right-hand member of the above inequality eventually becomes < 0, and hence

$$\leq -\frac{\nu(f_j'')}{\text{meas. } f_j''} \cdot \frac{\text{meas } f_j''}{\omega_j}.$$

Consequently, as j tends to infinity,

$$0 \leq \overline{\lim} \frac{n(f_j)}{\text{meas. } f_j} \leq -Ak'' + 0$$

which is a contradiction.

It follows that $\lim\ (i = \infty)\ \text{meas. } f_i''/\omega_i = 0$, and therefore that $\underline{\lim}\ \text{meas. } f'_i/\omega_i = k$. Consequently

$$(1.4) \qquad \lim_{i=\infty} \frac{\text{meas. } f_i''}{\text{meas. } f_i} = 0, \qquad \lim_{i=\infty} \frac{\text{meas. } f'_i}{\text{meas. } f_i} = 1.$$

From (1.2) we have

$$\frac{\nu(f'_i)}{\text{meas. } f'_i} \frac{\text{meas. } f'_i}{\text{meas. } f_i} \leq \frac{p(f_i)}{\text{meas. } f_i} < \frac{\nu(f'_i)}{\text{meas. } f'_i} \cdot \frac{\text{meas. } f'_i}{\text{meas. } f_i} + \frac{\epsilon_i}{\omega_i} \frac{\omega_i}{\text{meas. } f_i},$$

and from (1.4), letting i tend to infinity,

$$A \leq \underline{\lim}_{i=\infty} \frac{p(f_i)}{\text{meas. } f_i} \leq \overline{\lim} \frac{p(f_i)}{\text{meas. } f_i} \leq A.$$

Hence $p(r)$ has the unique derivative A. And from

$$\frac{n(r_i)}{\text{meas. } r_i} = \frac{p(r_i)}{\text{meas. } r_i} - \frac{\nu(r_i)}{\text{meas. } r_i}$$

it follows that $n(r)$ has the unique derivative zero.

A similar proof holds when $A < 0$, but if $A = 0$ the method must be slightly different, since both f'_i and f_i'' may now contain subsequences which are regular families with respect to the intervals ω_i. If f_i'' (f'_i) contains no regular family of parameter $> k/2$, with respect to the intervals ω_i, the entire sequence f'_i (f_i'') is a regular family with parameter $\geqq k/2$ with respect to these intervals, and the proof proceeds as before. On the other hand, if both f'_i and f_i'' contain regular families of this sort, f'_j and f_k'' respectively, we show directly that

$$\lim_{j=\infty} \frac{p(f_j)}{\text{meas.}\, f_j} = 0, \qquad \lim_{k=\infty} \frac{n(f_k)}{\text{meas.}\, f_k} = 0$$

from which, since $A = 0$,

$$\lim_{j=\infty} \frac{n(f_j)}{\text{meas.}\, f_j} = 0, \qquad \lim_{k=\infty} \frac{p(f_k)}{\text{meas.}\, f_k} = 0.$$

But the entire sequence $\{i\}$ is exhausted by the sequence $\{j\}$ and the sequence $\{k\}$, since for every i, either meas. f'_i/meas. f_i or meas. f_i''/meas. $f_i \geqq 1/2$. Hence for the entire sequence $\{i\}$

$$\lim_{j=\infty} \frac{p(f_i)}{\text{meas.}\, f_i} = \lim_{i=\infty} \frac{n(f_i)}{\text{meas.}\, f_i} = 0.$$

In order to prove that $t(r)$ has the unique derivative $|A|$ at $x = 0$, it is sufficient to remark that $t(r) = p(r) + n(r)$. Thus the proof of the lemma is complete.

In a Stieltjes integral, generalized or not, the function which appears in the integrand, before the differential sign, must be defined for every value of x in the interval to be considered. Accordingly we extend the definition of $\gamma(x)$ and write

(1.5) $$\gamma(x)/x]_{x=0} = \varliminf_{x=0} \gamma(x)/x,$$

admitting $+\infty$ as a possible value of this inferior limit.* With this extension, the following proposition is a consequence of Lemmas II and III.

LEMMA IV. *Let $g(x)$ be a function of bounded variation, $0 \leqq x \leqq a$, and $t(x)$ its total variation function. If $g(x)$ has a unique derivative at $x = 0$ (on regular families), namely $g'(0) = A$, then the generalized integrals*

$$\int_0^\delta [\gamma(x)/x]\, dg(x) = m_1(\delta), \quad \int_0^\delta [\gamma(x)/x]\, dt(x) = m_2(\delta), \quad 0 < \delta < a,$$

are convergent, and

---

* On account of the restriction imposed on $g(x)$ at $x = 0$, the truth of Lemma IV is independent of the definition of $\gamma(0)/0$, which may be arbitrary.

(1.6) $$\lim_{\delta=0} m_1(\delta) = 0, \qquad \lim_{\delta=0} m_2(\delta) = 0.$$

Since $t(x)$, by Lemma III, has a unique derivative at $x = 0$, it satisfies the conditions for $g(x)$, and it will be sufficient to prove the theorem for $g(x)$. We suppose first that $A = 0$, so that, choosing $g(0) = 0$,

$$g(x) = x\eta(x)$$

where $\eta(x)$ is of bounded variation in any closed interval (δ', δ) with $0 < \delta' < \delta$, and $|\eta(x)| < \eta_\delta$, $0 < x \leqq \delta$, with $\lim (\delta = 0)\ \eta_\delta = 0$. Without loss of generality we may assume $g(x)$ to be non-decreasing and $\eta(x)$ not negative, since if $g(x)$ has a zero derivative on regular families at $x = 0$, its positive and negative variation functions have also.

The generalized integral is given as the limit *

$$\lim_{n=\infty} \int_0^\delta \gamma_n(x)\, dg(x)$$

where $\gamma_n(x)$ is a sequence of functions, integrable with respect to $g(x)$, increasing with n, with $\lim\limits_{n=\infty} \gamma_n(x) = \gamma(x)/x$, $0 \leq x \leq \delta$, and such that for each n, $\int_0^\delta \gamma_n(x)\, dg(x)$ exists.

We take $n > 1/\delta$, N as the lower bound of $\gamma(x)/x$ in the interval $0 \leq x \leq 1/n$, and write

$$\begin{aligned} \gamma_n(x) &= N, & 0 \leq x \leq 1/n \\ \gamma_n(x) &= \gamma(x)/x, & 1/n < x \leq \delta. \end{aligned}$$

The function $\gamma_n(x)$ is bounded and measurable in the Borel sense, and is therefore integrable with respect to $g(x)$. For $x = 0$,

$$\lim_{n=\infty} \gamma_n(x) = \lim_{\underline{x=0}} \gamma(x)/x = \gamma(x)/x]_{x=0},$$

and for $x > 0$, $\gamma_n(x) = \gamma(x)/x$ for $n > 1/x$; hence $\lim\limits_{n=\infty} \gamma_n(x) = \gamma(x)/x$, $0 \leq x \leq \delta$. Moreover $\gamma_{n+k}(x) \geq \gamma_n(x)$ if $k > 0$. We have $N \leq n\gamma(1/n)$.

With these definitions we have

$$\begin{aligned} \int_0^\delta \gamma_n(x)\, dg(x) &= N \int_0^{1/n} dg(x) + \int_{1/n}^\delta \gamma_n(x)\, dg(x) \\ &= Ng(1/n) + \int_{1/n}^\delta \gamma_n(x)\, dg(x) \end{aligned}$$

* P. J. Daniell, *loc. cit.* (I), p. 289.

$$\leq Ng(1/n) + \int_{1/n}^{\delta} [\gamma(x)/x]\, dg(x).$$

But

$$Ng(1/n) \leq n\gamma(1/n) \cdot (1/n)\eta(1/n) < \eta_\delta \gamma(\delta)$$

and by Lemma II

$$\int_{1/n}^{x} \gamma(x)/x\, dg(x) = \int_{1/n}^{\delta} \gamma(x)\, d\eta(x) + \int_{1/n}^{\delta} [\gamma(x)/x]\, \eta(x)\, dx$$

$$\leq \left| \int_{1/n}^{\delta} \gamma(x)\, d\eta(x) \right| + m(\delta)\eta_\delta$$

where $m(\delta) = \int_0^\delta \gamma(x)/x\, dx$. Moreover, since $\gamma(x)$ is monotonic, an application of Lemma II to $\int_{1/n}^{\delta} d\,[\gamma(x)\eta(x)]$ yields the identity

$$\int_{1/n}^{\delta} \gamma(x)\, d\eta(x) = \gamma(\delta)\eta(\delta) - \gamma(1/n)\eta(1/n) - \int_{1/n}^{\delta} \eta(x)\, d\gamma(x),$$

and therefore the inequality

$$\left| \int_{1/n}^{\delta} \gamma(x)\, d\eta(x) \right| < 2\gamma(\delta)\eta_\delta + \eta_\delta\gamma(\delta) = 3\gamma(\delta)\eta_\delta.$$

Hence

$$\int_0^\delta \gamma_n(x)\, dg(x) < \eta_\delta\, [m(\delta) + 4\gamma(\delta)].$$

But the integral is an increasing function of n. Accordingly the integral of the Lemma exists; moreover from the inequality just written it approaches zero with δ. This is what was to be proved.

Suppose now that $A \neq 0$, and write

$$g(x) = xA + [g(x) - xA].$$

We know, as the property (A) of the general integral, that $\int_0^\delta [\gamma(x)/x]\, dg(x)$ exists and has the value

$$\int_0^\delta [\gamma(x)/x]\, dg(x) = \int_0^\delta [\gamma(x)/x]\, d(xA) + \int_0^\delta [\gamma(x)/x]\, d\,[g(x) - xA],$$

provided that both integrals of the second member of this equation exist. The existence of the second integral has just been demonstrated; it approaches zero with δ. The first integral of the right-hand member evidently can be shown to exist by the same sort of proof, since either xA or $-xA$ is a non-decreasing function. The integral is seen to have the value $Am(\delta)$, and also approaches zero with δ.

In fact, we have (taking for simplicity $A > 0$),

$$\int_0^\delta \gamma_n(x)\,d(xA) = N \cdot A \cdot 1/n + \int_{1/n}^\delta \gamma_n(x) A\,dx$$
$$= A\gamma(1/n) + A \int_{1/n}^\delta \gamma(x)/x\,dx$$
$$= A\gamma(1/n) + A\,[m(\delta) - m(1/n)]$$

which has the limit $Am(\delta)$ as n tends to ∞.

An obvious corollary of the theorem which has just been demonstrated is the following:

COROLLARY 1. *Let $\phi(x)$ be summable in the Lebesgue sense, in the interval $(0, a)$, and let it be the derivative of its integral at $x = 0$. Then the integral*

$$\int_0^\delta \phi(x)\gamma(x)/x\,dx$$

is convergent, as a Lebesgue integral, and approaches zero with δ.

Since $\gamma(kx)$, where k is a given positive constant, is a function $\bar{\gamma}(x)$, with the properties described for $\gamma(x)$, we may note the following fact.

COROLLARY 2. *Lemma IV remains valid if $\gamma(x)$ is replaced by $\gamma(kx)$, k const., > 0.*

We denote the corresponding $m_2(\delta)$ by $m_3(\delta)$.

2. *Convergence of the integrals* (A), (B) *on the boundary.* At the position when $P = Q$ the quotients $\cos(QP, n_P)/QP$, $\cos(QP, n_Q)/QP$ are not defined. In order to fix these quantities, we may for the purpose of the theorems of this note assign them any values whatever. We choose more or less arbitrarily the value zero.

If C has a unique tangent at Q and the condition (γ) is satisfied at Q, there will be a neighborhood of Q on C for which the projection of an element Δs of arc, on the tangent line, will have a measure as near Δs as we please. We take coördinates x, y with the origin at Q, and the x-axis as the direction of the curve in the positive sense, at Q; and choose δ small enough so that if $|x| \leq \delta$, $\gamma(s) \leq 1/2$. Let C_δ be the connected portion of C containing Q for which $|x_P| \leq \delta$. For convenience in reference we recall the following facts,* of which the verification is immediate.

For $P = (x, y)$ in C_δ,

* Evans and Miles, *loc. cit.*, p. 495.

$$
\begin{aligned}
&\cos(n_Q, n_P) \geq 1/2, \qquad ds \leq 2dx, \qquad s \leq 2x \\
&| dy | \leq \gamma(s)\, ds \leq 2\gamma(2 \,|\, x \,|\,) \,|\, dx \,|; \qquad | y | \leq 2 \,|\, x \,|\, \gamma(2 \,|\, x \,|\,) \\
(2.1)\quad &| \cos(QP, n_Q) | \leq | y | / | x | \leq 2\gamma(2 \,|\, x \,|\,) \\
&| \cos(QP, n_P) | \leq | \cos(QP, n_Q) | + | \cos(QP, n_P) - \cos(QP, n_Q) | \\
&\qquad \leq 2\gamma(2 \,|\, x \,|\,) + | \sphericalangle\, (n_Q, n_P) | \leq 3\gamma(2 \,|\, x \,|\,).
\end{aligned}
$$

By means of these inequalities and Lemma IV we can prove at once the following proposition.

LEMMA V. *Let the condition* (γ) *be satisfied at* Q, *and let* $\nu(e)$ *have at* Q *a unique derivative (on regular families),* $\nu'(Q) = A$, *with respect to* s. *Then the integrals*

$$(2.2)\quad \zeta = \int_C [\cos(QP, n_P)/QP]\, d\nu(e_P), \quad \zeta_1 = \int_C [\cos(QP, n_Q)/QP]\, d\nu(e_P)$$

are convergent, and

$$(2.3)\qquad \zeta = \lim_{\delta=0} \zeta'_\delta, \qquad \zeta_1 = \lim_{\delta=0} \zeta'_{1\delta}$$

where ζ'_δ, $\zeta'_{1\delta}$ *are the integrals* ζ, ζ_1 *except that they are extended merely over the portion* $C - C_\delta$ of C.

In fact if we let ζ_δ, $\zeta_{1\delta}$ be the corresponding integrals extended over C_δ, their integrands are dominated respectively by the functions $3\gamma(2 \,|\, x \,|\,)/|\, x \,|$ and $2\gamma(2 \,|\, x \,|\,)/|\, x \,|$; and if we let $t(e)$ be the total variation function of $\nu(e)$, we have the result that a dominant integral for both ζ_δ and $\zeta_{1\delta}$ is the following

$$\int_{C_\delta} [3\gamma(2x)/x]\, dt(e) = 3 \int_0^\delta [\gamma(2x)/x]\, dt_1(x) = 3m_3(\delta)$$

where $t_1(x) = [t(s_1) + t(s_2)]$, s_1 being the arc QP and s_2 the arc P_1Q such that $x_{P_1} = -x_P$, s_1 and s_2 being closed sets; $t_1(x)$ has a derivative at $x = 0$. But, since the integrands of ζ_δ and $\zeta_{1\delta}$ are both measurable in the Borel sense, and this dominant integral is convergent and approaches zero with δ, by Lemma IV, Corollary 2, the integrals ζ_δ and $\zeta_{1\delta}$ are both convergent and approach zero with δ.* Thus the theorem is proved.

The necessity of a smoothness condition at Q, of the nature of (γ), is made evident by the following example, in which C has a continuously turning tangent at Q, and $\nu(e)$ has a derivative (equal to zero) at Q, but where the integral ζ is not convergent.

---

* Daniell, *loc. cit.* (I), p. 290.

Take axes as before, and mark points $x = 1/2, 1/3, \cdots 1/k, \cdots$. At the point P_k $(x = 1/k)$, draw a segment, with a negative slope, such that its normal n_k satisfies the relation $\cos(x, n_k) = 1/(\log k)^{1/2}$. At the mid-point of the interval (P_{k+1}, P_k) draw a segment with a slope $1/k^2$. The broken line formed of these successive segments, with alternately positive and negative slopes, forms a portion of C abutting on Q. It is rectifiable, and

$$\lim\,(P = Q)\ |\ \sphericalangle\ (n_Q, n_P)\ | = 0,$$

provided that the vertices are smoothed, but the curve does not satisfy a (γ) condition at Q.

We form the function $\nu(e)$ by putting a mass a_k at each point P_k, and choose a_k so that $\nu(e)$ has a zero derivative, with respect to arc length, at Q. In fact, we take $a_k = [k(k+1)(\log k)^{1/2}]^{-1}$.

We have

$$\nu(s)/s \leq (x/s)(1/x) \sum_{k=r}^{\infty} a_k$$

where P_{r+1} is the P_k of lowest index contained in s. Hence

$$\begin{aligned}
\lim_{s=0}\,[\nu(s)/s] &\leq 1 \cdot \lim_{r=\infty}\,(r+1) \sum_{k=r}^{\infty} [k(k+1)(\log k)^{1/2}]^{-1} \\
&\leq \lim_{r=\infty}\,(r+1) \sum_{k=r}^{\infty} [1/k - 1/(k+1)]\ [1/(\log k)^{1/2}] \\
&\leq \lim_{r=\infty} \frac{r+1}{(\log r)^{1/2}} \left\{ \left(\frac{1}{r} - \frac{1}{r+1}\right) + \left(\frac{1}{r+1} - \frac{1}{r+2}\right) + \cdots \right\} \\
&\leq \lim\ [1/(\log r)^{1/2}] = 0.
\end{aligned}$$

Since $\nu(e)$ is a function of positive type it follows that the derivative on regular families is also zero.*

For the integral ζ, however, we have $\zeta = \zeta_\delta + \zeta'_\delta$

$$\begin{aligned}
\zeta_\delta = \int_Q^{Q_\delta} \frac{\cos(QP, n_P)}{QP}\ d\nu(e_P) &= \sum_{k=2}^{\infty} \frac{k}{(\log k)^{1/2}}\ a_k \\
&= \sum_{k=2}^{\infty} \frac{1}{(k+1)\log k},
\end{aligned}$$

which fails to be convergent.

In the example just given, the point charges may be replaced by uniform density distributions over intervals which become suitably short as x tends to

---

* De la Vallée Poussin, *Intégrales de Lebesgue, Fonctions d'ensemble, Classes de Baire*, Paris (1916), p. 60.

zero. In the following example the curve is still smoother and the mass distribution is uniform, but with a positive density N at $x = 0$. This time it is the integral ζ_1 which fails to be convergent.

Let $\lambda(x)$ be a continuous non-decreasing function of x, with $\lambda(0) = 0$, and such that $\int_0^\delta \lambda(x)/x\, dx$ fails to be convergent. Consider the curve C, symmetric about the y-axis, of which a portion near Q may be represented, for positive x, in the form $y = \int_0^x \lambda(x)\, dx$. Take $N > 0$ and let $\nu(e)$ be the function of sets N meas. $e = \nu(e)$, associated with the point function $g(P) = Ns$; it has the constant density N. We have, if the integral ζ_1 exists,

$$\zeta_{1\delta} = 2N \int_0^\delta [y/(x^2 + y^2)]\, [1 + \lambda(x)^2]^{1/2}\, dx > 2N \int_0^\delta [y/(x^2 + y^2)]\, dx$$

$$> N \int_0^\delta (y/x^2)\, dx, \text{ for } \delta \text{ small enough so that } x > y.$$

But by an integration by parts

$$\int_{\delta'}^\delta yx^{-2}\, dx = \Big[-(1/x) \int_0^x \lambda(x)\, dx \Big]_{\delta'}^{\delta} + \int_{\delta'}^\delta x^{-1} \lambda(x)\, dx,$$

and

$$\int_0^\delta yx^{-2}\, dx = -(1/\delta) \int_0^\delta \lambda(x)\, dx + \lim_{\delta=0} \int_{\delta'}^\delta x^{-1} \lambda(x)\, dx.$$

Hence the integral ζ_1 is not convergent.

The Fredholm method of treating the Dirichlet and Neumann problems is thus seen to be limited necessarily to curves which satisfy, at least almost everywhere, a condition of the type of (γ). The remark applies whether the boundary values are continuous or discontinuous. When the condition (γ) is not satisfied, in the former case many methods are still applicable; in the latter, the use of the Green's function leads to a general class of functions in which such problems are uniquely solvable,* but the connection with potentials of distributions of matter is lost, and the direct generalization to three dimensions is no longer feasible, on account of the lack of regularity of the Green's function.† The condition (γ) seems therefore to be a natural setting for the Fredholm method of treating the discontinuous boundary value problems in three dimensions.

---

* G. C. Evans, *The Logarithmic Potential Discontinuous Dirichlet and Neumann Problems*, New York (1927).

† J. J. Gergen, "Mapping of a General Type of Three Dimensional Region on a Sphere," *American Journal of Mathematics*, Vol. 52 (1930), pp. 197-224.

3. *Limiting values and the proof of Theorem I.* The determination of limiting values *in the wide sense* depends upon the following simple result. Let Q be the origin of coördinates, and $M = (x, y)$ a point such that $|y|/|x| \geq \tan q$ where q is a given constant, $0 < q < \pi/2$. Let P_1 be a point movable on the x-axis, of x coördinate x_1, and denote MP_1 by r_1. Then the maximum value of the ratio $QP_1/MP_1 = |x_1|/r_1$ is $\csc q$. In other words, this ratio is bounded as M approaches Q in the wide sense:

$$(3.1) \qquad |x_1|/r_1 \leq \csc q \quad \text{if} \quad |y|/|x| \geq \tan q.$$

Consider now the portion C_δ of C about Q, where the condition (γ) is assumed to hold. With axes as in § 2, let r denote MP, as before, and let P_1 be the projection of P on the x-axis. For P in C_δ, we have

$$\begin{aligned} |r - r_1| r_1^{-1} \leq |y_P| r_1^{-1} &< 2|x_1| \gamma(2|x_1|) r_1^{-1} \\ &< 2\gamma(2|x_1|) \csc q \leq 1/2 \end{aligned}$$

if δ is small enough. Consequently, $|x_1| r^{-1} = |x_1| r_1^{-1} [1 + (r - r_1)/r_1]^{-1} \leq 2|x_1| r_1^{-1}$ and

$$(3.2) \qquad |x_1|/r < 2 \csc q$$

for all M such that $|y|/|x| \geq \tan q$, provided that δ is taken small enough, depending on q but not on M.

We note also the following inequalities, in which the same notation is conserved, and ϕ_1 denotes the angle (r_1, n_Q). We have

$$\begin{aligned} |r^{-1} \sin \phi' - r_1^{-1} \sin \phi_1| &= |x_1 - x| \; |r^{-2} - r_1^{-2}| \\ &= r_1^{-2} r^{-2} |x_1 - x| \; (r + r_1) |r - r_1| \\ &\leq 2 |y_P| r^{-1} r_1^{-1} < \frac{2|x_1| \gamma(2|x_1|) x_1}{r_1 x_1 r} \\ &< 4 \csc^2 q \gamma(2|x_1|)/x_1. \end{aligned}$$

Similarly

$$\begin{aligned} |r^{-1} \cos \phi' - r_1^{-1} \cos \phi_1| &\leq \left| \frac{y_P - y}{r^2} + \frac{y}{r_1^2} \right| \\ &\leq \left| \frac{y_P}{r^2} \right| + \frac{|y| \; |r_1^2 - r^2|}{r^2 r_1^2} \\ &\leq 3 |y|_P \bar{r}^{-2}, \end{aligned}$$

where $\bar{r}$ is the smaller of the two values r, r_1. Hence

$$|r^{-1} \cos \phi' - r_1^{-1} \cos \phi_1| \leq 6 \left| \frac{x_1 x_1 \gamma(2|x_1|)}{\bar{r}\bar{r}|x_1|} \right| < 24 \csc^2 q \, \gamma(2|x_1|)/|x_1|.$$

2

But

$$r^{-1}\,|\sin\phi-\sin\phi'\,|\leq 2r^{-1}\left|\cos\frac{\phi+\phi'}{2}\sin\frac{\phi-\phi'}{2}\right|\leq r^{-1}\,|\,\phi-\phi'\,|$$

$$< 2\gamma(2\,|\,x_1\,|)\csc q/|\,x_1\,|$$

and

$$r^{-1}\,|\cos\phi-\cos\phi'\,|<2\gamma(2\,|\,x_1\,|)\csc q/|\,x_1\,|.$$

Consequently *there is a number $p>0$, such that for P in C_δ, we have, for δ small enough, depending on q but not on M,*

$$\begin{array}{lll} & |\,r^{-1}\sin\phi'-r_1^{-1}\sin\phi_1\,|, & |\,r^{-1}\cos\phi'-r_1^{-1}\cos\phi_1\,|, \\ (3.3) & |\,r^{-1}\sin\phi\;-r_1^{-1}\sin\phi_1\,|, & |\,r^{-1}\cos\phi\;-r_1^{-1}\cos\phi_1\,|, \\ & & \text{all} < p\,\gamma(2\,|\,x_1\,|)/x_1. \end{array}$$

If we denote any of these quantities by $f(x_1)$, we have

$$(3.4)\qquad\left|\int_{C_\delta}f(x_1)\,d\nu(e_P)\right|<pm_3(\delta),$$

by Lemma IV, Corollary 2.

On the basis of these inequalities consider the integral $u_2(M)$ in Theorem I (E_2), and denote by $u_{2\delta}$, $u'_{2\delta}$ the corresponding integrals over the portions C_δ and $C-C_\delta$ respectively of C. We write

$$\bar{u}_{2\delta}(M)=\int_{C_\delta}\frac{\sin\phi_1}{r_1}\,d\nu(e)=\int_{-\delta}^{\delta}\frac{\sin\phi_1}{r_1}\,d\nu(x_1)$$

with $\nu(x_1)=\nu(s_1)$, where s_1 is the portion of C_δ for which $x\leq x_1$.

We have, with M' the image of M, as in Theorem I, from (3.4),

$$|\,u_{2\delta}(M)-\bar{u}_{2\delta}(M)\,|<pm_3(\delta)$$
$$|\,u_{2\delta}(M')-\bar{u}_{2\delta}(M')\,|<pm_3(\delta)$$
$$\bar{u}_{2\delta}(M)-\bar{u}_{2\delta}(M')=0$$

so that, given ϵ, we have, independently of the position of M such that $|\,y/x\,|\geq\tan q$,

$$|\,u_{2\delta}(M)-u_{2\delta}(M')\,|<\epsilon/2$$

if δ is taken small enough so that $p<\epsilon/(4m_3(\delta))$.

We may now take M near enough to Q so that

$$|\,u'_{2\delta}(M)-u'_{2\delta}(M')\,|<\epsilon/2,$$

since these integrals are continuous as M passes through Q. Consequently, by taking M near enough to Q,

$$|\,u_2(M)-u_2(M')\,|<\epsilon.$$

But this is what was to be proved for (E_2) of Theorem I.

The proof of Theorem I (E_3), for $u_3(M)$, is identical with that for $u_2(M)$ just given. The examination of $u(M)$ and $u_1(M)$, however, requires a little more detail. Consider first the quantity

$$\bar{u}_{1\delta}(M) = \int_{C_\delta} \frac{\cos\phi_1}{r_1} d\nu(e) = \int_{C_\delta} \frac{\cos\phi_1}{r_1} d\nu(x_1)$$

where the function of bounded variation $\nu(x_1)$ has A for the value of its derivative at $x_1 = 0$. We have

$$\bar{u}_{1\delta}(M) = \int_{C_\delta} \frac{\cos\phi_1}{r_1} A\, dx_1 + \int_{C_\delta} \frac{\cos\phi_1}{r_1} d\{\nu(x_1) - Ax_1\}$$
$$= -A\theta(M,\delta) + \xi,$$

$\theta(M,\delta)$ being the angle subtended by the directed segment $-\delta \leq x_1 \leq \delta$ at M, and

$$|\xi| \leq \int_{-\delta}^{\delta} r_1^{-1} |\cos\phi_1|\, dt(x_1)$$

where $t(x_1)$ is a function of bounded variation with a zero derivative at $x_1 = 0$;

$$t(x_1) = x_1\eta(x_1)\,; \quad 0 \leq \eta(x_1) < \eta_\delta \quad \text{if} \quad -\delta \leq x_1 \leq \delta; \qquad \lim_{\delta=0} \eta_\delta = 0.$$

If $|x| \geq \delta$, $r_1^{-1}\cos\phi_1$ is monotonic in $-\delta \leq x \leq \delta$. Otherwise it is monotonic in each of the sub-intervals $(-\delta, x)$ and (x, δ). We write

$$|\xi| \leq \left| \int_{-\delta}^{x} r_1^{-1}\cos\phi_1\, dt(x_1) \right| + \left| \int_{x}^{\delta} r_1^{-1}\cos\phi_1\, dt(x_1) \right|$$

$$\int_x^\delta \frac{\cos\phi_1}{r_1} dt(x_1) = t(x_1)\frac{\cos\phi_1}{r_1}\Big]_x^\delta - \int_x^\delta x_1\eta(x_1)\, d\,\frac{\cos\phi_1}{r_1}.$$

Now

$$\left| t(x_1)\frac{\cos\phi_1}{r_1}\Big]_x^\delta \right| = \left| \frac{x_1}{r_1}\eta(x_1)\cos\phi_1 \Big]_x^\delta \right| \leq 2\eta_\delta \csc q$$

$$\left| \int_x^\delta x_1\eta(x_1)\, d\left(\frac{\cos\phi_1}{r_1}\right) \right| < \eta_\delta \left| \int_x^\delta |x_1|\, d\,\frac{\cos\phi_1}{r_1} \right| = \eta_\delta \left| \; |x_1|\,\frac{\cos\phi_1}{r_1}\Big]_x^\delta - \int_x^\delta \frac{\cos\phi_1}{r_1} d\,|x_1| \right|$$
$$< 2\eta_\delta \csc q + \eta_\delta\,\pi/2.$$

Consequently, whether $|x| \gtrless \delta$,

$$|\xi| \leq \eta_\delta\,\{8 \csc q + \pi\}$$

and

(3.5) $\quad |\bar{u}_{1\delta}(M) + A\theta(M,\delta)| \leq p_1\eta_\delta, \qquad p_1$ const., depending on q.

But from (3.3) and Lemma V,

$$\begin{aligned} | u_\delta(M) - \bar{u}_{1\delta}(M) | &< p m_3(\delta) \\ \zeta_\delta(M) \qquad\qquad &< 3 m_3(\delta). \end{aligned}$$

Hence, given ϵ, we can take δ, small enough, depending on q but not on M; provided that $| y/x | \geq \tan q$, so that

$$\begin{aligned} | u_\delta(M) + A\theta(M,\delta) | &\leq \epsilon/3 \\ | \zeta_\delta | &\leq \epsilon/3. \end{aligned}$$

And since $u'_\delta(M)$ is continuous as M passes through Q, we can take M near enough to Q so that

$$| u(M) - u_\delta(M) - (\zeta - \zeta_\delta) | \leq \epsilon/3.$$

By taking M near enough to Q, therefore, with $| y/x | \geq \tan q$, we have

$$| u(M) + A\theta(M,\delta) - \zeta | \leq \epsilon.$$

Similarly

$$| u_1(M) + A\theta(M,\delta) - \zeta | \leq \epsilon.$$

Since, however, no matter what the value of δ,

$$\lim_{M=\pm Q} \theta(M,\delta) = \pm \pi,$$

the above inequalities imply the relations (E), (E_1) of Theorem I. This completes the proof of the theorem.

4. *Potentials in the neighborhood of a vertex.* We may speak of a Q of C as a *spiral* point, if one of the four quantities $\bar{\theta}_+(Q)$, $\bar{\theta}_-(Q)$, $\underline{\theta}_+(Q)$, $\underline{\theta}_-(Q)$, defined in the Introduction, as we approach Q from one direction or the other on C, is infinite. We say that Q, on the other hand, is a vertex of C, if at Q the forward and backward directions of the curve exist in the ordinary sense, and if these directions are not the same and do not differ by π; in the former case Q is a *cusp* point. We denote the directions of the forward and backward tangents at Q by t_+ and t_- respectively. From these the directions of the interior "normals" n_{Q+} and n_{Q-} are established. The positive angle α from the forward tangent t_+ to the backward tangent t_- is called the angle at the vertex.

We may say that the curve satisfies a condition (γ_α) at Q if the angle at Q is α and if, with the function $\gamma(s)$ as already defined,

(4.1) $\quad | \sphericalangle (n_{Q\pm}, n_P) | < \gamma(s), \quad P$ on C in $\pm$ sense from Q.

The function $\nu(e)$ has a unique forward (backward) derivative on regular families at Q if it has no point value at Q, and if the limit is unique for families of sets e_i on the forward (backward) branch of the curve from Q, regular with respect to the smallest intervals ω_i which contain Q and e_i.

With these definitions we may state the proposition about the convergence of the integral $u(M)$, when $M = Q$, a vertex.

LEMMA V. *Let the condition* (γ_a) *be satisfied at* Q, *and* $\nu(e)$ *have unique forward and backward derivatives,* A_+ *and* A_- *respectively, (on regular families) at* Q. *Then the integral* ζ *converges, and*

$$\zeta = \lim_{\delta=0} \zeta'_\delta.$$

By C_δ we may understand the portion of C in the neighborhood of Q whose projections on t_+ and t_- have measures δ_+ and δ_- respectively, both $< \delta$.

Similarly (3.3) applies, where the angle q is the lower bound of the value of the smallest absolute angle which QM makes with the directions t_+ or t_-, and n_{Q+} or n_{Q-} is used according as P is forward or backward from Q.

We are thus led to a useful theorem. Let M in T_+ or in T_- approach Q in such a way that QM approaches a limiting direction which makes an angle β with t_+, the angle β being positive, $0 < \beta < \alpha$, if M is in T_+ and negative, $\alpha - 2\pi < \beta < 0$, if M is in T_-.

THEOREM II. *If* C *satisfies* (γ_a) *at* Q *and* $\nu(e)$ *has unique forward and backward derivatives,* A_+ *and* A_- *respectively (on regular families) at* Q, *then*

$$\lim_{M=Q\pm} u(M) = (\beta \mp \pi)A_+ + (\alpha - \beta - \pi)A_- + \zeta. \tag{4.2}$$

A special case of this theorem is the following corollary:

COROLLARY. *If* $A_+ = A_- = A$

$$\lim_{M=Q+} u(M) = (\alpha - 2\pi)A + \zeta, \qquad \lim_{M=Q-} u(M) = \alpha A + \zeta, \tag{4.3}$$

and the limits exist in the wide sense.

5. *Extensions and applications.* The methods used in this paper permit the extension of the results without serious difficulty to three dimensions. Thus in both two and three dimensions they constitute a natural medium for the setting of the Fredholm method. In fact, in the case where the (γ) condition is satisfied uniformly for all points Q of C and $\nu(e)$ is absolutely continuous—which includes as a special case the continuous boundary value conditions—the Fredholm theory may be established more systematically and just as simply as for boundaries of class C''.

There is also an immediate application to the Cauchy integral extended over the boundary C. If this integral is written in terms of a boundary function $f+ig$, summable (L) on C, for points M in T_+,

$$w(z)=\frac{1}{2\pi i}\int_C \frac{f(t)+ig(t)}{t-z}\,dt,$$

a necessary and sufficient condition that $w(z)$ take on the values $f(t)+ig(t)$ in the wide sense, almost everywhere, or in fact, on a set of positive measure, on C, is that $w(z)$ vanish identically for z outside C. This result depends on the uniqueness theorem of Riesz and Lusin and Priwaloff,* and is established immediately by writing $w(z)$ in terms of integrals of the form $u(M)$ and $u_2(M)$, $M=z$, and applying Theorem I above. In obtaining this result there is thus involved an extension of the method of Fichtenholz † for the corresponding theorem in the case of the circle. Nevertheless, the latter method uses the inverse point and the special formulae related to the circle, and so is not directly applicable. Likewise, whereas in the case of the circle the class of functions involved is obtained by explicit formulae, here, results must be obtained from general considerations of the class of functions given by integrals like $u(M)$, $u_2(M)-u_2(M')$. If the (γ) condition is uniform, this class may be determined.

These results will be demonstrated in later notes.

THE RICE INSTITUTE,
HOUSTON, TEXAS.

* N. Luzin and J. Priwaloff, *Annales scientifiques de l'École Normale Supérieure*, Vol. 41 (1925), p. 159.

† G. Fichtenholz, *Fundamenta Mathematicae*, Vol. 13 (1929), pp. 1-33 (see p. 23).

GENERALIZATIONS TO HIGHER DIMENSIONS OF THE CAUCHY INTEGRAL FORMULA.

By D. G. Fulton and G. Y. Rainich.

In the first part of the paper formulas are presented which are immediate generalizations to higher dimensions of the Cauchy integral formula giving the value of an analytic function in a point surrounded by a contour or closed hypersurface in terms of its value on the contour or hypersurface. Then in connection with Volterra's theory of conjugate functions, about which a brief outline is given, formulas are written which are extensions of a much more general nature of the Cauchy integral formula.

Immediate Generalizations to Higher Dimensions.

1. The generalizations may be better approached from an historical point of view. We may start with Newton's Law of Gravitation which says that the components of the gravitational field produced by a single material point are proportional to

$$x/r^3, \quad y/r^3, \quad z/r^3.$$

It can be readily verified that such a field (and also a sum of any number of such fields) satisfies the differential equations

$$X_1 + Y_2 + Z_3 = 0, \quad X_2 = Y_1, \quad X_3 = Z_1, \quad Y_3 = Z_2 \tag{1.1}$$

where X_1 means $\partial X/\partial x$, X_2 means $\partial X/\partial y$, etc. For this reason we call the set of equations (1.1) the Newtonian differential equations. It is clear that by setting $Z = 0$ we obtain a special case in which X and Y are functions of x and y alone, and satisfy the equations

$$X_1 + Y_2 = 0, \quad X_2 = Y_1$$

which may be interpreted as the Cauchy-Riemann equations for the function

$$W = Y + iX.$$

Thus we see that the Newtonian equations are a generalization of the Cauchy-Riemann equations. Later we shall take up still more general cases; at present we want to derive a formula which bears the same relationship to the Newtonian equations as the Cauchy integral formula bears to the Cauchy-Riemann equations.

2. We begin by rewriting the Cauchy integral formula

$$f(t) = \frac{1}{2\pi i}\int \frac{f(z)\,dz}{z-t}$$

in a form which readily yields itself to generalization. Making $t = 0$ we have

$$f(0) = \frac{1}{2\pi i}\int \frac{f(z)\,dz}{z}\ .$$

We write $f(z) = W = Y + iX$, $z = x + iy$, $dz = dx + idy$ and $-i(dx + idy) = (\alpha + \beta i)\,ds$, where α and β are the direction cosines of the normal to the contour, and the formula becomes

$$2\pi X_0 = \int\left(X\,\frac{\alpha x + \beta y}{r^2} + Y\,\frac{\beta x - \alpha y}{r^2}\right) ds.$$

$$2\pi Y_0 = \int\left(X\,\frac{\alpha x + \beta y}{r^2} + Y\,\frac{\alpha x + \beta y}{r^2}\right) ds.$$

Once expressed in this manner it is easy to prove the formula. We shall give the proof here because the proof in higher cases will be modelled on this one.

The proof consists of the following two steps.

a). The integrals vanish on a contour not surrounding the origin because by the divergence theorem

$$\int [X(\alpha x + \beta y)/r^2 + Y(\beta x - \alpha y)/r^2]\,ds$$
$$= \int \{(\partial/\partial x)[(Xx - Yy)/r^2] + (\partial/\partial y)[(Xy + Yx)/r^2]\}\,dA$$

and the integrand of the last integral equals

$$(x/r^2)(\partial X/\partial x + \partial Y/\partial y) + (y/r^2)(\partial X/\partial y - \partial Y/\partial x)$$
$$+ X\,[\partial(x/r^2)/\partial x + \partial(y/r^2)/\partial y] + Y\,[\partial(x/r^2)/\partial y - \partial(y/r^2)/\partial x]$$

which vanishes since each parenthesis is zero.

b). It follows that the value of the integral is the same for all contours around the origin; we may integrate, for instance, around a circle of radius r. We have in this case $x/r = \alpha$, $y/r = \beta$, which makes the first integral become $\int X(ds/r)$ and we know that it is independent of r, whence it is easily seen that its value is $2\pi X_0$.

3. We can now generalize the formula to the Newtonian case. By analogy we write:

$$4\pi X_0 = \int\left(X\,\frac{\alpha x + \beta y + \gamma z}{r^3} + Y\,\frac{\beta x - \alpha y}{r^3} + Z\,\frac{\gamma x - \alpha z}{r^3}\right) d\omega$$

$$4\pi Y_0 = \int \left(X \, \frac{\alpha y - \beta x}{r^3} + Y \, \frac{\alpha x + \beta y + \gamma z}{r^3} + Z \, \frac{\gamma y - \beta z}{r^3} \right) d\omega$$

$$4\pi Z_0 = \int \left(X \, \frac{\alpha z - \gamma x}{r^3} + Y \, \frac{\beta z - \gamma y}{r^3} + Z \, \frac{\alpha x + \beta y + \gamma z}{r^3} \right) d\omega.$$

The proof of this formula is exactly similar to that in two dimensions.

The integral formula in two dimensions may be considered as a limit of this one in three dimensions.

4. The three integrals above may be expressed very easily in one integral. Let

$$W = iX + jY + kZ, \qquad n = i\alpha + j\beta + k\gamma$$

and

$$M = i\,[\partial(1/r)/\partial x] + j\,[\partial(1/r)/\partial y] + k\,[\partial(1/r)/\partial z].$$

Then we have

$$4\pi W_0 = \int W \cdot n \cdot M \cdot d\omega.$$

The product under the integral sign is to be understood as the product of quaternions.*

5. The generalization to higher dimensions is obvious. For four dimensions, for example, we consider four functions X, Y, Z, T and set up as analogous to the Newtonian equations the system

$$X_1 + Y_2 + Z_3 + T_4 = 0$$

$$X_2 = Y_1, \quad X_3 = Z_1, \quad X_4 = T_1, \quad Y_3 = Z_2, \quad Y_4 = T_2, \quad Z_4 = T_3.$$

In this case we get in the same fashion,

$$2\pi^2 X_0 = \int \left(X \, \frac{\alpha x + \beta y + \gamma z + \delta t}{r^4} + Y \, \frac{\beta x - \alpha y}{r^4} + Z \, \frac{\gamma x - \alpha y}{r^4} + T \, \frac{\delta x - \alpha t}{r^4} \right) d\tau \dagger \text{ etc.}$$

The proof again does not present any difficulty.

6. For n dimensions we consider n functions F_i, and we have the system of equations ‡

---

* A. C. Dixon, *Quarterly Journal*, Vol. 35 (1904), pp. 283-296. Also D. Iwanenko and K. Nikolski, *Zeitschrift für Physik*, Vol. 63 (1930), p. 129.

† For the measure of a hypersphere in n dimensions see Somerville, "An Introduction to the Geometry of n Dimensions."

‡ Here and in what follows a Greek letter used as an index means summation with respect to that index.

$$\partial F_\sigma/\partial x_\sigma = 0, \quad \partial F_i/\partial x_j = \partial F_j/\partial x_i$$
$$(\sigma = \textstyle\sum 1, 2, 3 \cdots n;\ i, j = 1, 2, 3 \cdots n).$$

The corresponding integral formula is

$$\frac{2\pi^{\frac{1}{2}n}}{\Gamma(\frac{1}{2}n)}(F_i)_0 = \int \left(F_i \frac{\alpha_\sigma x_\sigma}{r^n} + F_\sigma \frac{\alpha_\sigma x_i - \alpha_i x_\sigma}{r^n} \right) d\tau.$$

This may be considered as the immediate generalizations of the Cauchy integral formula but it is possible to apply the same principles to a still more general case.

More General Extensions to Higher Dimensions.

7. In 1889, Volterra * founded the theory of conjugate functions of which all the preceding cases are special cases but which contain many others. We adopt the integral point of view in order to show in what this generalization consists. The systems of functions F_i which we considered so far give zero integrals on contours and closed hypersurfaces, that is, for spaces of n dimensions on closed manifolds of 1 and $n - 1$ dimensions. Volterra's generalization consists in considering systems which give zero integrals on closed manifolds of r and $n - r$ dimensions.

Without going into details (which the reader may look up in a paper by F. D. Murnaghan in the *Physical Review* (1921), pp. 73-87 and in his book *Vector Analysis and the Theory of Relativity*, pp. 72-77), we mention the following points:

The functions in this case will be $F_{i_1 i_2 \ldots i_r}$, which change the sign if two indices are interchanged, and consequently vanish when two indices are identical. These functions give rise to another system of functions $R_{i_1 i_2 \ldots i_{n-r}}$ by means of the formulas

$$F_{i_1 i_2 \ldots i_r} = R_{i_{r+1} i_{r+2} \ldots i_n},$$

where the n subscripts form an even permutation of the numbers $1, 2, 3, \cdots n$.

The system lends itself to integration over an r-dimensional surface and the integral will be

$$\int F_{\rho_1 \rho_2 \ldots \rho_r} \alpha_{\rho_1 \rho_2 \ldots \rho_r} \tag{7.1}$$

where

---

* V. Volterra, "Sulle funzioni conjugate," *Rendiconti dei Lincei* (4), Vol. 5 (1889), pp. 599-611. For developments analogous to those presented here see Volterra's paper in *Rendiconti di Palermo*, Vol. 3 (1889), pp. 260-272.

$$\alpha_{i_1 i_2 \ldots i_r} = \begin{vmatrix} dx_{i_1}, & dx_{i_2}, & \cdots, & dx_{i_r} \\ \delta_1 x_{i_1}, & \delta_2 x_{i_2}, & \cdots, & \delta_1 x_{i_r} \\ \delta_2 x_{i_1}, & \delta_2 x_{i_2}, & \cdots, & \delta_2 x_{i_r} \\ \cdot & \cdot & \cdot & \cdot \\ \delta_{r-1} x_{i_1}, & \delta_{r-1} x_{i_2}, & \cdots, & \delta_{r-1} x_{i_r} \end{vmatrix}.$$

On an $n-r$ manifold the integral will be of the form

$$(7.2) \qquad \int R_{\rho_1 \rho_2 \ldots \rho_{n-r}} \alpha_{\rho_1 \rho_2 \ldots \rho_{n-r}}$$

where $\alpha_{i_1 i_2 \ldots i_{n-r}}$ is similar to $\alpha_{i_1 i_2 \ldots i_r}$. Using the divergence theorem

$$\int_V \frac{\partial F_{i_1 i_2 \ldots i_{r-1} i \rho}}{\partial x_\rho} \, d\tau = \int_S F_{i_1 i_2 \ldots i_{r-1} i \rho} \alpha_{i_1 i_2 \ldots i_{r-1} i \rho}$$

and an analogous relation for the R's, we find the following system of differential equations to be equivalent to the vanishing of the integrals (7.1) and (7.2):

$$(7.3) \qquad \frac{\partial F_{i_1 i_2 \ldots i_{r-1} i \rho}}{\partial x_\rho} = 0,$$

$$(7.4) \qquad \frac{\partial R_{i_1 i_2 \ldots i_{n-r-1} i \rho}}{\partial x_\rho} = 0.$$

Making use of the relationship between the F's and the R's we may write (7.4) as

$$(7.5) \qquad (-1)^\rho \frac{\partial F_{i_1 i_2 \ldots i'\rho \ldots i_{r+1}}}{\partial x_\rho} = 0$$

where a primed index means that all the indices appear with the exception of the one primed.

8. It is clear that these equations are generalizations of the Cauchy-Riemann and the Newtonian equations. Besides they contain many other cases of which the simplest is that corresponding to $n = 4$, $r = 2$. The set of differential equations giving the vanishing of surface integrals is

$$\partial F_{ia}/\partial x_a = 0, \qquad \partial R_{ia}/\partial x_a = 0.$$

These equations differ only in sign before certain terms from the fundamental equations of electro-dynamics in free space in Maxwell's theory. This difference in sign is essential for our purposes since it would substitute for integrals over a closed surface integrals over a surface extending to infinity—compare for the case of Maxwell's equations with the article by Murnaghan referred to above.

We derived these equations as conditions for the vanishing of *surface*

integrals, but they also may be shown (by applying the reasoning used at the end of section 7) to be conditions for the vanishing of certain *hypersurface* integrals, namely, the eight integrals

$$\int F_{i\tau}\alpha_\tau d\omega \quad \text{and} \quad \int R_{i\tau}\alpha_\tau d\omega \qquad (i = 1, 2, 3, 4).$$

We may expect that for these equations the integral formula will be analogous to the Cauchy formula; in fact, we find

$$2\pi^2(F_{ij})_0 = \int\left(F_{ij}\frac{\alpha_\tau x_\tau}{r^4} + F_{\tau j}\frac{\alpha_\tau x_i - \alpha_i x_\tau}{r^4} + F_{i\tau}\frac{\alpha_\tau \alpha_j - \alpha_j x_\tau}{r^4}\right) d\omega$$

where α_i $(i = 1, 2, 3, 4)$ are the direction cosines of the normal to the hypersurface.

9. In the case where n is arbitrary and $r = 3$ we may write the integral in a form similar to the ones for $r = 2$. We get

$$\frac{2\pi^{\frac{1}{2}n}}{\Gamma(\frac{1}{2}n)}(F_{ijk})_0 = \int\Big(F_{ijk}\frac{\alpha_\tau x_\tau}{r^n} + F_{\tau jk}\frac{\alpha_\tau x_i - \alpha_i x_\tau}{r^n}$$
$$+ F_{i\tau k}\frac{\alpha_\tau x_j - \alpha_j x_\tau}{r^n} + F_{ij\tau}\frac{\alpha_\tau x_k - \alpha_k x_\tau}{r^n}\Big) d\omega.$$

The proof of this formula consists of two steps as in the two-dimensional case.

(9.1) The value of the integral is independent of the hypersurface surrounding the origin. To prove this we show that

$$\sum_{l=1}^{n}\frac{\partial(\text{coef. of } \alpha_l)}{\partial x_l} = 0.$$

Coefficient of α_l,

$l = i$: $F_{\tau jk}(x_\tau/r^n)$
$l = j$: $F_{i\tau k}(x_\tau/r^n)$
$l = k$: $F_{ij\tau}(x_\tau/r^n)$
$l \neq ijk$: $F_{ijk}(x_l/r^n) - F_{ljk}(x_i/r^n) - F_{ilk}(x_j/r^n) - F_{ijl}(x_k/r^n)$

(a) $$\sum_{l=1}^{n}\frac{\partial(\text{coef. of } \alpha_l)}{\partial x_l}$$
$$= \sum_{l=1}^{n}\left(\frac{x_l}{r^n}\frac{\partial F_{ijk}}{\partial x_l} - \frac{x_i}{r^n}\frac{\partial F_{ljk}}{\partial x_l} - \frac{x_j}{r^n}\frac{\partial F_{ilk}}{\partial x_l} - \frac{x_k}{r^n}\frac{\partial F_{ijl}}{\partial x_l}\right) \qquad (l \neq i, j, k).$$

(b)+ $$(x_\tau/r^n)(\partial F_{\tau jk}/\partial x_i + \partial F_{i\tau k}/\partial x_j + \partial F_{ij\tau}/\partial x_k)$$

(c)+ $$\sum_{l=1}^{n}(F_{ijk}/r^n) + 3F_{ijk}/r^n \qquad (l \neq i, j, k).$$

(d)+ $$\sum_{l=1}^{n}[\partial(1/r^n)/\partial x_l]\,\{F_{ijk}x_l - x_iF_{ljk} - x_jF_{ilk} - x_kF_{ijl}\} \qquad (l \neq i, j, k).$$

(e)[+] $[\partial(1/r^n)/\partial x_i]\, x_\tau F_{\tau jk} + [\partial(1/r^n)/\partial x_j]\, x_\tau F_{i\tau k} + [\partial(1/r^n)/\partial x_k]\, x_\tau F_{ij\tau}$.

Lines (a) and (b) combined vanish as a result of the differential equations

$$\partial F_{i_1 i_2 i_\rho}/\partial x_\rho = 0,$$

$$(-1)^\rho \frac{\partial F_{i_1 i_2 \ldots\, i'_\rho \ldots\, i_i}}{\partial x_\rho} = 0.$$

In (d) and (e) combined ($l \neq i, j, k$ in the next five formulas)

$$\sum_{l=1}^{n} [\partial(1/r^n)/\partial x_l]\, x_l F_{ijk} = -(n/r^{n+2}) F_{ijk} \sum_{l=1}^{n} x_l^2 \;;$$

$$-x_i \sum_{l=1}^{n} \{[\partial(1/r^n)/\partial x_l]\, F_{ljk}\} + [\partial(1/r^n)/\partial x_i]\, x_\tau F_{\tau jk} = -n(x_i^2/r^{n+2}) F_{ijk};$$

$$-x_j \sum_{l=1}^{n} \{[\partial(1/r^n)/\partial x_l]\, F_{ilk}\} + [\partial(1/r^n)/\partial x_j]\, x_\tau F_{i\tau k} = -n(x_j^2/r^{n+2}) F_{ijk};$$

$$-x_k \sum_{l=1}^{n} \{[\partial(1/r^n)/\partial x_l]\, F_{ijl}\} + [\partial(1/r^n)/\partial x_k]\, x_\tau F_{ij\tau} = -n(x_k^2/r^{n+2}) F_{ijk}.$$

Hence (c), (d) and (e) together give

$$\sum_{l=1}^{n} (F_{ijk}/r^n) + 3F_{ijk}/r^n - (nF_{ijk}/r^{n+2})\left(\sum_{l=1}^{n} x_l^2 + x_i^2 + x_j^2 + x_k^2\right) = 0.$$

Therefore

$$\sum_{l=1}^{n} \frac{\partial(\text{coef. of } \alpha_l)}{\partial x_l} = 0.$$

(9.2) Since it is true that the value of the integral is the same on all hypersurfaces we may integrate around a hypersphere in n dimensions. Here $x_i/r = \alpha_i$, and the integrand reduces to F_{ijk}/r^{n-1}. Since the integral is independent of r we easily obtain as its value

$$\frac{2\pi^{\frac{1}{2}n}}{\Gamma(\frac{1}{2}n)} (F_{ijk})_0.$$

The same reasoning would apply to the most general case; namely, when both n and r are arbitrary, $r < n$. The integral formula in this case is

$$\frac{2\pi^{\frac{1}{2}n}}{\Gamma(\frac{1}{2}n)} (F_{i_1 i_2 \ldots i_r})_0 = \int \Bigg(F_{i_1 i_2 \ldots i_r} \frac{\alpha_\tau x_\tau}{r^n} + F_{\tau i_2 i_3 \ldots i_r} \frac{\alpha_\tau x_{i_1} - \alpha_{i_1} x_\tau}{r^n}$$
$$+ F_{i_1 \tau i_3 \ldots i_r} \frac{\alpha_\tau x_{i_2} - \alpha_{i_2} x_\tau}{r^n} + \cdots + F_{i_1 i_2 \ldots i_{r-1}\tau} \frac{\alpha_\tau x_{i_r} - \alpha_{i_r} x_\tau}{r^n} \Bigg) d\omega.$$

This is the general formula we wanted to obtain. All the preceding formulas are special cases of it.

FOURIER DEVELOPMENTS FOR CERTAIN PSEUDO-PERIODIC FUNCTIONS IN TWO VARIABLES.

By Miguel A. Basoco.

1. *Introduction.* Jacobi,* in his investigations on the dynamics of a rotating rigid body, was led to the discovery of certain pseudo-periodic functions which, using Hermite's † nomenclature, belong to the class of doubly periodic functions of the second kind. Expressed in terms of the Jacobi theta functions, they have the form

$$\phi_{\alpha\beta\gamma}(x, y) \equiv \vartheta'_1 \frac{\vartheta_\alpha(x+y)}{\vartheta_\beta(x)\vartheta_\gamma(y)}, \tag{1}$$

where x and y are independent complex variables, $\vartheta'_1 = \vartheta_0\vartheta_2\vartheta_3$, and (α, β, γ) are restricted to certain sixteen triads ‡ out of the possible sixty-four which can be selected from the numbers 0, 1, 2, 3. Hermite and Kronecker § obtained the Fourier series developments of these functions; recently, these have been recalculated and corrected by Bell,¶ who has shown that they play

* Jacobi, *Werke*, Bd. 2, pp. 291-351.

† Hermite, "Sur quelques applications des fonctions elliptiques," *Comptes Rendus*, Vol. 85 (1877), . . ., Vol. 94 (1882); *Oeuvres*, Vol. 3, p. 267; "Sur une application de la théorie des fonctions doublement périodiques de seconde espèce," *Annales de l'École Normale Supérieure* (3), Vol. 2 (1885), p. 303; *Oeuvres*, Vol. 4, pp. 190 and 199-200.

‡ The series for the functions corresponding to the remaining forty-eight triads have not been calculated; the method due to Teixeira [see footnote †, p. 243] indicates why only sixteen appear in a natural manner; the remaining can not be obtained by his method. However, by multiplying each of the known $\phi_{\alpha\beta\gamma}(x, y)$ by a suitable Jacobian elliptic function, it is possible to obtain expansions for the forty-eight missing functions, *altered*, however, by a multiplicative factor of the second degree in the theta constants ϑ_a, $a = 0, 2, 3$. From a different point of view, these functions have recently been studied by D. A. F. Robinson in a University of Chicago doctoral dissertation (1930). His results are to be published in a forthcoming number of the *Proceedings of the Royal Society of Canada, Toronto.* The series obtained by these methods have a more complex structure than those corresponding to the $\phi_{\alpha\beta\gamma}(x, y)$ as defined in equation (1).

§ Kronecker, *Monatsberichte der Berliner Akademie* (1881), pp. 1165 to 1172; *Werke*, Bd. IV, pp. 309 to 318. See also Tannery-Molk, *Fonctions Elliptiques*, Vol. 3, pp. 120-131.

¶ E. T. Bell, "Arithmetical Paraphrases, II," *Transactions of the American Mathematical Society*, Vol. 22, No. 2 (1921), p. 207; *Colloquium Publications*, Vol. 7, p. 88. See also, *Messenger of Mathematics*, Vol. 49 (Sept., 1919), p. 83.

an all-important rôle in certain investigations in the theory of numbers, as for example, in establishing some of the very general formulae involving numerical functions which were published by Liouville * without proof or indication as to their source.

A systematic study of the doubly periodic functions of the second kind, as a class, has been made by F. Gomes Teixeira.† He has given a general theory for obtaining the trigonometric expansions of these functions and has indicated, by an example, how it may be applied to the functions $\phi_{\alpha\beta\gamma}(x, y)$. He takes, as a starting point, the function of x, $\vartheta_\alpha(x+y)/\vartheta_\beta(x)$, regarding y temporarily as a parameter; his analysis then leads in a direct manner to the series for the functions (1).

In the present paper we obtain the arithmetized Fourier developments of certain functions $\Theta_{\alpha\beta\gamma}(x, y)$, which are related to the $\phi_{\alpha\beta\gamma}(x, y)$ by the identity

$$(2) \qquad \Theta_{\alpha\beta\gamma}(x, y) \equiv \frac{\vartheta'_1}{\vartheta_\beta(x)} \cdot \phi_{\alpha\beta\gamma}(x, y) = \vartheta'^2_1 \frac{\vartheta_\alpha(x+y)}{\vartheta_\beta^2(x)\vartheta_\gamma(y)} .$$

In as much as the arithmetized series for the indicated factors in (2) are known,‡ this identity may be used to obtain arithmetical paraphrases of the general Liouville type. These, as well as other such formulae derivable from the $\Theta_{\alpha\beta\gamma}(x, y)$ will involve *incomplete* numerical functions in *two variables.* So far as the writer is aware no expansions involving incomplete numerical functions in *more than one variable* have appeared in the literature. The importance of such functions may be judged from a recent paper by Uspensky,§ who by purely arithmetic methods has established certain formulae of the paraphrase type which have reference to incomplete functions in three variables.

As suggested by Teixeira's example, mentioned above, we begin with $\vartheta_\alpha(x+y)/\vartheta_\beta^2(x)$ regarding it as a function of x only; since, as will be evident from what follows, these are *not* doubly periodic of the second kind,

---

* Liouville, "Sur quelques formules générales qui peuvent être utiles dans la théorie des nombres," Eighteen memoirs in *Journal de Mathématiques* (1858-1865).

† Teixeira, "Sur le developpement des fonctions doublement périodiques de seconde espèce en serie trigonometrique," *Crelle's Journal für Mathematik*, Bd. 125 (1901), pp. 301-318.

‡ E. T. Bell, "Theta Expansions Useful in Arithmetic," *Messenger of Mathematics*, Vol. 54 (1924).

§ J. V. Uspensky, *Bulletin of the American Mathematical Society*, Vol. 36 (October, 1930), p. 743; *American Journal of Mathematics*, Vol. 50 (1928), No. 1. See also an article by A. Oppenheim in *Quarterly Journal of Mathematics*, Oxford series, Vol. 2, no. 7, (Sept., 1931).

Teixeira's method is not applicable and hence our analysis will run along somewhat different lines. The method * which we shall use is not limited in its application to the particular functions here treated, but may also be applied to functions of the form $\vartheta_\alpha(x+y)/[\vartheta_\beta(x)]^r$, where r is an integer greater than unity. For $r > 2$, however, the results are fairly complex.†

As in the case of the $\phi_{\alpha\beta\gamma}(x,y)$, our method leads to sixteen functions $\Theta_{\alpha\beta\gamma}(x,y)$; from these, sixteen more may be obtained by permuting the variables x and y. The remaining ninety-six may be obtained from these thirty-two by multiplying them by suitable elliptic functions. In this connection, see note (‡), page 242.

2. *Group I.* The functions $\Theta_{\alpha\beta\gamma}(x,y)$ appear in four groups of four each, such that those within a given group may be derived from any one in the same group by replacing the variables (x,y) in turn by $(x+\pi/2, y)$, $(x, y+\pi/2)$ and $(x+\pi/2,\ y+\pi/2)$. To obtain the expansions for the functions in the first group, we may begin with

$$f(x)=\frac{\vartheta_1(x+y)}{\vartheta_1{}^2(x)}, \tag{3}$$

which, as follows from the properties of the theta functions,‡ satisfies the relations

$$\begin{cases} f(x+\pi)=-f(x)\,;\ f(x+2\pi)=f(x) \\ f(x+n\pi\tau)=(-1)^n q^{n^2} e^{2ni(x-y)} f(x), \qquad [i=(-1)^{1/2}]. \end{cases} \tag{4}$$

where, $q=e^{i\pi\tau}$, $0<\operatorname{amp}\tau<\pi$, so that $|q|<1$, and n is any positive or negative integer, including zero. The Laurent expansion of $f(x)$ relative to the pole $x=0$ is readily seen to be

$$A_2/x^2+A_1/x+P(x), \tag{5}$$

in which $A_2=\vartheta_1(y)/\vartheta'_1{}^2$ and $A_1=\vartheta'_1(y)/\vartheta'_1{}^2$.

Consider a parallelogram $ABCD$ in the x-complex plane, composed of $n+1$ double cells (of width 2π) above and n below the real axis. In this parallelogram, let the poles of the function in the r-th double cell have affixes $a_r=r\pi\tau$ and $b_r=\pi+r\pi\tau$. From (4) and (5) it follows that in the neighborhood of these poles, $f(x)$ is represented by the series

---

* See Whittaker and Watson, ***Modern Analysis***, Fourth Edition, p. 139.

† The method is theoretically applicable to functions of the form

$$[\vartheta_\alpha(x+y)]^s/[\vartheta_\beta(x)]^r,$$

provided $s<r$.

‡ We use the small theta notation of Jacobi, with his ϑ replaced by ϑ_0, see *Werke*, I, s. 501. In this notation the argument of the circular functions does not, as in some others, contain the factor π.

$$(6)\quad \begin{cases} (-1)^r q^{r^2} e^{-2irv} \left(\dfrac{A_2}{(x-a_r)^2} + \dfrac{A_1 + 2irA_2}{x-a_r} + \cdots \right), \\ -(-1)^r q^{r^2} e^{-2irv} \left(\dfrac{A_2}{(x-b_r)^2} + \dfrac{A_1 + 2irA_2}{x-b_r} + \cdots \right), \end{cases}$$

respectively.

We now consider the auxiliary function

$$(7)\qquad \Phi(t) = \frac{f(t)}{\sin(t-x)}.$$

Within the parallelogram $ABCD$, this function has poles at $t = x$, $t = \pi + x$, $t = a_r$, $t = b_r$, $r = 0, \pm 1, \pm 2, \cdots, \pm n$. (Whenever any of the poles lie on the boundary of the parallelogram, as is the case here, it is permissible to derange the mesh formed by the period cells so that all poles may lie within.) The residues relative to the first two poles are the same and equal to $f(x)$; from (6) and (7) it may be readily seen that the residues relative to a_r and b_r are, in both cases,

$$(-1)^r q^{r^2} e^{-2irv}[(A_1 + 2irA_2)\operatorname{cosec}(a_r - x) - A_2 \operatorname{cosec}(a_r - x)\operatorname{ctn}(a_r - x)].$$

On applying Cauchy's theorem to the function $\Phi(t)$, with $ABCD$ as contour of integration and then allowing n to become infinite we obtain equation (8) below. Clearly, the integral may be replaced by four other integrals, two of which will cancel due to the periodicity of $\Phi(t)$, while the other two will tend to zero as n tends to infinity, due to the presence of the factor q^{n^2}. Obviously, it is the existence of such a factor as this, which makes this method of expansion possible. We may, thus, write

$$(8)\quad \begin{cases} \vartheta_1'^2 \dfrac{\vartheta_1(x+y)}{\vartheta_1^2(x)\vartheta_1(y)} = \dfrac{\vartheta_1'(y)}{\vartheta_1(y)} \cdot \displaystyle\sum_{r=-\infty}^{\infty} (-1)^r q^{r^2} e^{-2irv} \operatorname{cosec}(x - r\pi\tau) \\ \qquad + 2i \displaystyle\sum_{r=-\infty}^{\infty} (-1)^r q^{r^2} r e^{-2irv} \operatorname{cosec}(x - r\pi\tau) \\ \qquad + \displaystyle\sum_{r=-\infty}^{\infty} (-1)^r q^{r^2} e^{-2irv} \operatorname{cosec}(x - r\pi\tau)\operatorname{ctn}(x - r\pi\tau). \end{cases}$$

This expansion is valid for all values of x and y with the exception of those for which the function is not defined. For arithmetical purposes, the expansion, as just obtained, is not useful; it is necessary to transform it into a Fourier series. This may be accomplished by replacing each sum in (8) by two others in which the indices of summation range from $r = 1$ to $r = \infty$, pairing off corresponding terms and making use of the following developments:

$$(9)\qquad \operatorname{cosec}(x+r\pi\tau) = -2i\sum_{m=1}^{\infty} q^{mr}e^{mix}, \qquad (m=1,3,5,7,\cdots),$$

$$(10)\qquad \operatorname{cosec}(x-r\pi\tau) = 2i\sum_{m=1}^{\infty} q^{mr}e^{-mix},$$

$$(11)\qquad \operatorname{cosec}(x+r\pi\tau)\operatorname{ctn}(x+r\pi\tau) = -2\sum_{m=1}^{\infty} mq^{mr}e^{mix},$$

$$(12)\qquad \operatorname{cosec}(x-r\pi\tau)\operatorname{ctn}(x-r\pi\tau) = -2\sum_{m=1}^{\infty} mq^{mr}e^{-mix}.$$

On making the necessary reductions and applying the remarks made at the beginning of this section we obtain the expansions corresponding to the functions in the first group. They are valid for all values of y for which the functions are defined and for all values of x such that

$$-I(\pi\tau) < I(x) < I(\pi\tau),$$

i. e., such that their representative points in the x-plane lie within a strip whose boundaries are straight lines parallel to the axis of reals and passing through the points $x=\pi\tau$ and $x=-\pi\tau$.

With regard to notation we point out that the finite sums which appear as the coefficients of q^r, refer to all the integer divisors d, δ or t, τ of r subject to whatever restrictions may be indicated. We shall let n be an unrestricted integer greater than zero, m an odd integer >0, α and β positive integers of the form $4k+1$ and $4k+3$, respectively; the divisor τ shall always be odd and positive.

Group I.

$$(2n=d\delta,\ d<(2n)^{1/2},\ \delta-d\equiv 1,\ \operatorname{mod} 2.)$$

$$(I_a):\left\{\begin{aligned} \Theta_{111}(x,y) &= \vartheta_1'^2\frac{\vartheta_1(x+y)}{\vartheta_1^2(x)\vartheta_1(y)} = \frac{\cos x}{\sin^2 x} - 4\sum q^{2n} \\ &\quad \Big(\sum(-1)^d(d+\delta)\cos((\delta-d)x+2dy)\Big) \\ &\quad + \frac{\vartheta_1'(y)}{\vartheta_1(y)}\left(\frac{1}{\sin x} + 4\sum q^{2n}\Big(\sum(-1)^d\sin((\delta-d)x+2dy)\Big)\right) \end{aligned}\right.$$

$$(I_b):\left\{\begin{aligned} \Theta_{221}(x,y) &= \vartheta_1'^2\frac{\vartheta_2(x+y)}{\vartheta_2^2(x)\vartheta_1(y)} = -\frac{\sin x}{\cos^2 x} + 4\sum q^{2n} \\ &\quad \Big(\sum(-1)^{(d+\delta-1)/2}(d+\delta)\sin((\delta-d)x+2dy)\Big) \\ &\quad + \frac{\vartheta_1'(y)}{\vartheta_1(y)}\left(\frac{1}{\cos x} + 4\sum q^{2n}\Big(\sum(-1)^{(d+\delta-1)/2}\cos((\delta-d)x+2 \right. \end{aligned}\right.$$

$$(I_c): \begin{cases} \Theta_{212}(x,y) = \vartheta_1'^2 \dfrac{\vartheta_2(x+y)}{\vartheta_1^2(x)\vartheta_2(y)} = \dfrac{\cos x}{\sin^2 x} - 4\sum q^{2n} \\ \qquad \Big(\sum (d+\delta)\cos((\delta-d)x+2dy)\Big) \\ \quad + \dfrac{\vartheta_2'(y)}{\vartheta_2(y)}\Big(\dfrac{1}{\sin x} + 4\sum q^{2n}\Big(\sum \sin((\delta-d)x+2dy)\Big)\Big). \end{cases}$$

$$(I_d): \begin{cases} \Theta_{122}(x,y) = \vartheta_1'^2 \dfrac{\vartheta_1(x+y)}{\vartheta_2^2(x)\vartheta_2(y)} = \dfrac{\sin x}{\cos^2 x} - 4\sum q^{2n} \\ \qquad \Big(\sum (-1)^{(\delta-d-1)/2}(d+\delta)\sin((\delta-d)x+2dy)\Big) \\ \quad - \dfrac{\vartheta_2'(y)}{\vartheta_2(y)}\Big(\dfrac{1}{\cos x} + 4\sum q^{2n}\Big(\sum (-1)^{(\delta-d-1)/2}\cos((\delta-d)x+2dy)\Big)\Big). \end{cases}$$

3. *Group II.* The details involved in the calculation of the expansions for the functions of this group are entirely similar to those in the preceding section and therefore will be omitted. We find that if we begin with $f(x) = \vartheta_2(x+y)/\vartheta_0^2(x)$ we are led to the series

$$(13) \quad \begin{cases} \vartheta_1'^2 \dfrac{\vartheta_2(x+y)}{\vartheta_0^2(x)\vartheta_3(y)} = -\dfrac{\vartheta_3'(y)}{\vartheta_3(y)} \sum\limits_{r=-\infty}^{\infty} q^{(2r+1)^2/4} e^{-i(2r+1)y} \operatorname{cosec}(x-(2r+1)\pi\tau/2) \\ \quad - i \sum\limits_{r=-\infty}^{\infty} q^{(2r+1)^2/4}(2r+1)e^{-i(2r+1)y} \operatorname{cosec}(x-(2r+1)\pi\tau/2) \\ \quad - \sum\limits_{r=-\infty}^{\infty} q^{(2r+1)^2/4} e^{-i(2r+1)y} \operatorname{cosec}(x-(2r+1)\pi\tau/2)\operatorname{ctn}(x-(2r+1)\pi\tau/2) \end{cases}$$

With the aid of the series similar to (9)—(12), the preceding expansion may be reduced to a Fourier series. Thus we get the following group, in which,

$$-\tfrac{1}{2} I(\pi\tau) < I(x) < \tfrac{1}{2} I(\pi\tau).$$

Group II.

$(\beta = t\tau,\ \beta \equiv 3 \bmod 4,\ \tau < \beta^{1/2}.)$

$$(II_a): \begin{cases} \Theta_{203}(x,y) = \vartheta_1'^2 \dfrac{\vartheta_2(x+y)}{\vartheta_0^2(x)\vartheta_3(y)} \\ \quad = 2\sum q^{\beta/4}\Big(\sum (t+\tau)\cos\Big(\dfrac{t-\tau}{2}x+\tau y\Big) \\ \qquad - 4\,\dfrac{\vartheta_3'(y)}{\vartheta_3(y)} \sum q^{\beta/4}\Big(\sum \sin\dfrac{t-\tau}{2}x+\tau y\Big)\Big). \end{cases}$$

$$(II_b): \begin{cases} \Theta_{133}(x,y) = \vartheta_1'^2 \dfrac{\vartheta_1(x+y)}{\vartheta_3^2(x)\vartheta_3(y)} \\ \qquad = 2\sum q^{\beta/4}\left(\sum(-1)^{(t-\tau-2)/4}(t+\tau)\sin\left(\dfrac{t-\tau}{2}x+\tau y\right)\right) \\ \qquad + 4\dfrac{\vartheta_3'(y)}{\vartheta_3(y)}\sum q^{\beta/4}\left(\sum(-1)^{(t-\tau-2)/4}\cos\left(\dfrac{t-\tau}{2}x+\tau y\right)\right) \end{cases}$$

$$(II_d): \begin{cases} \Theta_{100}(x,y) = \vartheta_1'^2 \dfrac{\vartheta_1(x+y)}{\vartheta_0^2(x)\vartheta_0(y)} \\ \qquad = 2\sum q^{\beta/4}\left(\sum(-1)^{(\tau-1)/2}(t+\tau)\sin\left(\dfrac{t-\tau}{2}x+\tau y\right)\right) \\ \qquad + 4\dfrac{\vartheta_0'(y)}{\vartheta_0(y)}\sum q^{\beta/4}\left(\sum(-1)^{(\tau-1)/2}\cos\left(\dfrac{t-\tau}{2}x+\tau y\right)\right). \end{cases}$$

$$(II_e): \begin{cases} \Theta_{230}(x,y) = \vartheta_1'^2 \dfrac{\vartheta_2(x+y)}{\vartheta_3^2(x)\vartheta_0(y)} \\ \qquad = -2\sum q^{\beta/4}\left(\sum(-1)^{(t+\tau)/4}(t+\tau)\cos\left(\dfrac{t-\tau}{2}x+\tau y\right)\right) \\ \qquad + 4\dfrac{\vartheta_0'(y)}{\vartheta_0(y)}\sum q^{\beta/4}\left(\sum(-1)^{(t+\tau)/4}\sin\left(\dfrac{t-\tau}{2}x+\tau y\right)\right). \end{cases}$$

4. *Group III.* The functions of the first two groups have the period 2π while those in the next two have the period π. Hence, the appropriate auxiliary function $\Phi(t)$ to be used in connection with these last groups is

$$(14) \qquad \Phi(t) = \frac{f(t)}{\tan(t-x)},$$

in place of (7). Also, the contour of integration is modified so that the width of the parallelogram $ABCD$ in § 2, is π, i. e., the width of a period cell. Proceeding as before, if we let $f(x)$ be $\vartheta_0(x+y)/\vartheta_0^2(x)$, we are led to the series

$$(15) \quad \begin{cases} \vartheta_1'^2 \dfrac{\vartheta_0(x+y)}{\vartheta_0^2(x)\vartheta_1(y)} \\ = i\dfrac{\vartheta_1'(y)}{\vartheta_1(y)}\sum\limits_{r=-\infty}^{\infty}(-1)^{r+1}q^{(2r+1)^2/4}e^{-i(2r+1)y}\operatorname{ctn}(x-(2r+1)\pi\tau/2) \\ + \sum\limits_{r=-\infty}^{\infty}(-1)^{r}q^{(2r+1)^2/4}(2r+1)e^{-i(2r+1)y}\operatorname{ctn}(x-(2r+1)\pi\tau/2) \\ + i\sum\limits_{r=-\infty}^{\infty}(-1)^{r+1}q^{(2r+1)^2/4}e^{-i(2r+1)y}\operatorname{cosec}^2(x-(2r+1)\pi\tau/2). \end{cases}$$

The reduction of this expression to a Fourier series can be carried out with the aid of the following expansions:

(16) $$\operatorname{ctn}(x - r\pi\tau) = i\left[1 + 2\sum_{n=1}^{\infty} q^{2rn} e^{-2inx}\right],$$

(17) $$\operatorname{ctn}(x + r\pi\tau) = -i\left[1 + 2\sum_{n=1}^{\infty} q^{2rn} e^{2inx}\right],$$

(18) $$\operatorname{cosec}^2(x - r\pi\tau) = -4\sum_{n=1}^{\infty} nq^{2rn} e^{-2inx},$$

(19) $$\operatorname{cosec}^2(x + r\pi\tau) = -4\sum_{n=1}^{\infty} nq^{2rn} e^{2nix}.$$

In the following series the variable x is restricted so that

$$-\tfrac{1}{2} I(\pi\tau) < I(x) < \tfrac{1}{2} I(\pi\tau).$$

Group III.

$(\alpha = t\tau,\ \alpha \equiv 1 \bmod 4,\ \tau < \alpha^{1/2};\ m \equiv 1 \bmod 2,\ (-1/m) = (-1)^{(m-1)/2}.)$

$$(III_a):\left\{\begin{aligned}\Theta_{001}(x,y) = {}& \vartheta_1'^2 \frac{\vartheta_0(x+y)}{\vartheta_0^2(x)\vartheta_1(y)} - 2\sum(-1/m)\,mq^{m^2/4}\sin my \\ & + 2\sum q^{a/4}\left(\sum(-1/\tau)(t+\tau)\sin\left(\frac{t-\tau}{2}x+\tau y\right)\right) \\ & + \frac{\vartheta_1'(y)}{\vartheta_1(y)}\left\{2\sum(-1/m)\,q^{m^2/4}\cos my\right. \\ & \left. + 4\sum q^{a/4}\left(\sum(-1/\tau)\cos\left(\frac{t-\tau}{2}x+\tau y\right)\right)\right\}.\end{aligned}\right.$$

$$(III_b):\left\{\begin{aligned}\Theta_{331}(x,y) = {}& \vartheta_1'^2 \frac{\vartheta_3(x+y)}{\vartheta_3^2(x)\vartheta_1(y)} - 2\sum(-1/m)\,mq^{m^2/4}\sin my \\ & + 2\sum q^{a/4}\left(\sum(-1)^{(t+\tau-2)/4}(t+\tau)\sin\left(\frac{t-\tau}{2}x+\tau y\right)\right) \\ & + \frac{\vartheta_1'(y)}{\vartheta_1(y)}\left\{2\sum(-1/m)\,q^{m^2/4}\cos my\right. \\ & \left. + 4\sum q^{a/4}\left(\sum(-1)^{(t+\tau-2)/4}\cos\left(\frac{t-\tau}{2}x+\tau y\right)\right)\right\}.\end{aligned}\right.$$

$$(III_c):\left\{\begin{aligned}\Theta_{302}(x,y) = {}& \vartheta_1'^2 \frac{\vartheta_3(x+y)}{\vartheta_0^2(x)\vartheta_2(y)} - 2\sum q^{m^2/4} m\cos my \\ & + 2\sum q^{a/4}\left(\sum(t+\tau)\cos\left(\frac{t-\tau}{2}x+\tau y\right)\right) \\ & - \frac{\vartheta_2'(y)}{\vartheta_2(y)}\left\{2\sum q^{m^2/4}\sin my\right. \\ & \left. + 4\sum q^{a/4}\left(\sum\sin\left(\frac{t-\tau}{2}x+\tau y\right)\right)\right\}.\end{aligned}\right.$$

$$(III_d): \begin{cases} \Theta_{032}(x,y) = \vartheta'^2_1 \dfrac{\vartheta_0(x+y)}{\vartheta_3^2(x)\vartheta_2(y)} = 2\sum q^{m^2/4} m \cos my \\ \quad + 2\sum q^{a/4}\left(\sum (-1)^{(t-\tau)/4}(t+\tau)\cos\left(\dfrac{t-\tau}{2}x+\tau y\right)\right) \\ \quad - \dfrac{\vartheta'_2(y)}{\vartheta_2(y)}\left\{ 2\sum q^{m^2/4}\sin my \right. \\ \quad \left. + 4\sum q^{a/4}\left(\sum (-1)^{(t-\tau)/4}\sin\left(\dfrac{t-\tau}{2}x+\tau y\right)\right)\right\}. \end{cases}$$

5. *Group IV.* If we let $f(x)$ be $\vartheta_0(x+y)/\vartheta_1^2(x)$ and proceed as in the preceding section we are led to the expansion

$$(20) \quad \begin{cases} \vartheta'^2_1 \dfrac{\vartheta_0(x+y)}{\vartheta_1^2(x)\vartheta_0(y)} = 2i\sum_{r=-\infty}^{\infty}(-1)^r r q^{r^2} e^{-2iry}\operatorname{ctn}(x-r\pi\tau) \\ \quad + \sum_{r=-\infty}^{\infty}(-1)^r q^{r^2} e^{-2iry}\operatorname{cosec}^2(x-r\pi\tau) \\ \quad + \dfrac{\vartheta'_0(y)}{\vartheta_0(y)}\sum_{r=-\infty}^{\infty}(-1)^r q^{r^2} e^{-2iry}\operatorname{ctn}(x-r\pi\tau). \end{cases}$$

Using (16) to (19) in connection with (20) we obtain the expansions in Group IV. Here the variable x is restricted by the condition

$$-I(\pi\tau) < I(x) < I(\pi\tau).$$

Group IV.

$(n = d\delta,\ \delta - d \equiv 0 \bmod 2,\ d < n^{1/2}.)$

$$(IV_a): \begin{cases} \Theta_{010}(x,y) = \vartheta'^2_1 \dfrac{\vartheta_0(x+y)}{\vartheta_1^2(x)\vartheta_0(y)} = \dfrac{1}{\sin^2 x} - 4\sum_{n=1}^{\infty}(-1)^n n q^{n^2}\cos 2ny \\ \quad - 4\sum q^n\left(\sum (-1)^d(d+\delta)\cos((\delta-d)x+2dy)\right) \\ \quad + \dfrac{\vartheta'_0(y)}{\vartheta_0(y)}\left\{\operatorname{ctn} x + 2\sum_{n=1}^{\infty}(-1)^n q^{n^2}\sin 2ny \right. \\ \quad \left. + 4\sum q^n\left(\sum (-1)^d \sin((\delta-d)x+2dy)\right)\right\}. \end{cases}$$

$$(IV_b):\quad\left\{\begin{aligned}&\Theta_{320}(x,y)=\vartheta'^2_1\frac{\vartheta_3(x+y)}{\vartheta_2^2(x)\vartheta_0(y)}=\frac{1}{\cos^2 x}-4\sum(-1)^n nq^{n^2}\cos 2ny\\&\quad-4\sum q^n\Big(\sum(-1)^{(d+\delta)/2}(d+\delta)\cos\left((\delta-d)x+2dy\right)\Big)\\&\quad-\frac{\vartheta'_0(y)}{\vartheta_0(y)}\Big\{\tan x-2\sum_{n=1}^{\infty}(-1)^n q^{n^2}\sin 2ny\\&\quad-4\sum q^n\Big(\sum(-1)^{(d+\delta)/2}\sin\left((\delta-d)x+2dy\right)\Big)\Big\}.\end{aligned}\right.$$

$$(IV_c):\quad\left\{\begin{aligned}&\Theta_{313}(x,y)=\vartheta'^2_1\frac{\vartheta_3(x+y)}{\vartheta_1^2(x)\vartheta_3(y)}=\frac{1}{\sin^2 x}-4\sum_{n=1}^{\infty}nq^{n^2}\cos 2ny\\&\quad-4\sum q^n\Big(\sum(d+\delta)\cos\left((\delta-d)x+2dy\right)\Big)\\&\quad+\frac{\vartheta'_3(y)}{\vartheta_3(y)}\Big\{\operatorname{ctn} x+2\sum_{n=1}^{\infty}q^{n^2}\sin 2ny\\&\quad+4\sum q^n\Big(\sum\sin\left((\delta-d)x+2dy\right)\Big)\Big\}.\end{aligned}\right.$$

$$(IV_d):\quad\left\{\begin{aligned}&\Theta_{023}(x,y)=\vartheta'^2_1\frac{\vartheta_0(x+y)}{\vartheta_2^2(x)\vartheta_3(y)}=\frac{1}{\cos^2 x}-4\sum_{n=1}^{\infty}nq^{n^2}\cos 2ny\\&\quad-4\sum q^n\Big(\sum(-1)^{(\delta-d)/2}(d+\delta)\cos\left((\delta-d)x+2dy\right)\Big)\\&\quad-\frac{\vartheta'_3(y)}{\vartheta_3(y)}\Big\{\tan x-2\sum q^{n^2}\sin 2ny\\&\quad-4\sum q^n\Big(\sum(-1)^{(\delta-d)/2}\sin\left((\delta-d)x+2dy\right)\Big)\Big\}.\end{aligned}\right.$$

6. In the preceding expansions, some of the terms are multiplied by expressions of the form $\vartheta'_a(y)/\vartheta_a(y)$; the arithmetized series for these are well known * and we list them here for the sake of completeness.

$$\frac{\vartheta'_0(y)}{\vartheta_0(y)}=4\sum q^n\left(\sum\sin 2ty\right),\qquad (n=t\tau), \tag{21}$$

$$\frac{\vartheta'_1(y)}{\vartheta_1(y)}=\operatorname{ctn} y+4\sum q^{2n}\left(\sum\sin 2dy\right),\qquad (n=d\delta), \tag{22}$$

* See, for example Bell's "Theta Expansions Useful in Arithmetic," *loc. cit.*, p. 169. Also, certain reduction formulae in his paper in the *Giornale di Matematiche*, Vol. 61 (1923), pp. 213-228, will be found useful.

$$(23) \qquad \frac{\vartheta'_2(y)}{\vartheta_2(y)} = -\tan y + 4\sum q^{2n}\left(\sum(-1)^d \sin 2dy\right), \qquad (n = d\delta),$$

$$(24) \qquad \frac{\vartheta'_3(y)}{\vartheta_3(y)} = 4\sum(-1)^n q^n\left(\sum \sin 2ty\right), \qquad (n = t\tau).$$

Expansions (21) and (24) are valid for $-\frac{1}{2}I(\pi\tau) < I(y) < \frac{1}{2}I(\pi\tau)$, while (22) and (23) are valid for $-I(\pi\tau) < I(y) < I(\pi\tau)$.

When these are introduced into our expansions the coefficients of the general power of q, will then be a finite sum ranging over the solutions of certain *quaternary quadratic forms.*

7. In conclusion we may point out that our formulae will yield expansions for certain doubly periodic functions of the *third kind.* Thus, if we put $x = y$ in, say, (II_0) and apply the transformation of the second order, we obtain an expansion for $\vartheta'_1[\vartheta_1(x)\vartheta_2(x)\vartheta_3(x)/\vartheta_0{}^2(x)]$. If we let $x = -y$, we obtain series for functions of the form $\vartheta'_1{}^2\vartheta_2/\vartheta_3{}^2(x)\vartheta_0(x)$. Again, if we put $y = 0$, wherever it is permissible, as in (II_0), we get series for functions of the type $\vartheta'_1\vartheta_2\vartheta_3[\vartheta_1(x)/\vartheta_0{}^2(x)]$. With regard to the last two cases just mentioned, it should, perhaps, be stated that the method of the present paper, when applied to functions involving a single variable, does not differ in essence from the method given by Appell's theory * of the decomposition of doubly periodic functions of the third kind with more poles than zeros. For, as is shown by the writer elsewhere,† it is possible, by the method here indicated, to derive Appell's fundamental function and his formula of decomposition for this type of function.

Finally, the writer wishes to acknowledge his indebtedness to Professor E. T. Bell for his suggestion to the effect that the functions here treated might lead, when expanded into the series here given, to incomplete functions in more than one variable.

THE UNIVERSITY OF NEBRASKA.

---

* Appell, *Annales Scientifiques de l'École Normale Supérieure,* Vols. 1, 2, 3, 5, series 3, 1884-1888.

† *Acta Mathematica,* Vol. 57 (1931), pp. 95-100.

SOLUTION OF PARTIAL DIFFERENTIAL EQUATIONS BY MEANS OF HYPERVARIABLES.

By P. W. KETCHUM.

1. *Introduction.* The simplicity and elegance of the fact that the analytic functions of an ordinary complex variable constitute a general solution of Laplace's equation in two variables, have induced many attempts to obtain a similar general solution of Laplace's equation in three or more variables. B. Peirce * described a hypervariable which he thought would succeed. But Bechk-Widmanstetter † proved that no hypervariable of three units can be used. Whittaker ‡ by generalizing in another direction succeeded in obtaining a general solution by means of an integral. His work was extended by Bateman.§ An interesting general solution of Laplace's equation in four variables was obtained by Nicolesco.¶

The writer ‖ has obtained a hypervariable associated with an algebra containing an infinity of units such that the totality of analytic functions (*i. e.*, as in ordinary complex variables, functions expressible as power series or functions possessing unique derivatives) constitutes a general solution of Laplace's equation in three dimensions. The hypervariable is defined as follows:

$$w = ze_{-1} + xe_0 + ye_1$$

where e_{-1}, e_0, e_1 are three of an infinite set of units whose multiplication table may be defined by putting

$$e_0 = 1, \quad e_{-n} = i^n \sin n\alpha, \quad e_n = i^n \cos n\alpha,$$

$$[\text{e. g.}, e_{-4} \cdot e_3 = i^7 \sin 4\alpha \cdot \cos 3\alpha = \tfrac{1}{2} i^7 \sin 7\alpha - \tfrac{1}{2} i \sin \alpha = \tfrac{1}{2} e_{-7} - \tfrac{1}{2} e_{-1}.]$$

It should be noticed that although w contains only three units, functions of w will in general contain more than three units. For example:

* *Linear Associative Algebra*, p. 124.

† *Monatshefte für Mathematik und Physik*, Vol. 23 (1912), p. 257.

‡ *Mathematische Annalen*, Vol. 57 (1902), p. 333; Whittaker and Watson, *Modern Analysis*, p. 388.

§ *Proceedings of the London Mathematical Society*, Ser. 2, Vol. 1 (1904), p. 451.

¶ *Comptes Rendus*, Vol. 185 (1927), p. 442. See also Hedrick and Ingold, *Transactions of the American Mathematical Society*, Vol. 27 (1925), p. 551.

‖ *American Journal of Mathematics*, Vol. 51 (1929), p. 179. The first equation in Section 3 should read $f(w) = \Sigma f_i(x, y, z) e_i$. Similar equations in the middle of p. 182 and at the bottom of p. 183 should read $f(w) = \Sigma f_i e_i$.

$$w^2 = \tfrac{1}{2}(2x^2 - y^2 - z^2)e_0 + 2xye_1 + 2xze_{-1} + \tfrac{1}{2}(y^2 - z^2)e_2 + yze_{-2}.$$

Functions other than polynomials will in general contain an infinity of units. In all cases the coefficients of the units satisfy Laplace's equation.

The formal relation between this solution and Whittaker's is somewhat like the relation between Fourier series and Fourier integrals, but the viewpoints are quite different. In the two dimensional case the real and imaginary parts of powers of a complex variable form the circular harmonics. In an analogous way, if

$$w^m = \sum_{j=-m}^{m} f_j(x, y, z)e_j,$$

the functions f_j are proportional to the solid spherical harmonics. For example, in the above expression for w^2 the coefficients of the e's are the solid spherical harmonics of order 2 except for constant factors. In fact this is perhaps the most natural and direct approach to the subject of spherical harmonics. Formally this result is already known and has been used as the basis of elementary expositions of the subject.*

In the present paper it is intended to make a further and more general study of the relations between linear partial differential equations with constant coefficients and hypervariables. In sections 3 and 4 it will be shown that to every such equation there is associated a linear associative and commutative algebra. Some properties of these algebras will be determined. In section 5 a hypervariable is found which gives a general solution of Airy's equation. Section 6 gives the hypervariables for Laplace's equation in n variables, including the wave equation as a special case. Section 7 contains an application to simultaneous partial differential equations.

2. *Form of the general solution of a differential equation.* Consider the homogeneous partial differential equation with constant coefficients:

$$P(\partial/\partial x_1, \partial/\partial x_2, \cdots \partial/\partial x_n) \cdot U = 0$$

where P is a homogeneous polynomial of degree r and with either real or ordinary complex coefficients. (Any equation with constant coefficients in n variables may be written as a homogeneous equation in $n+1$ variables.)†

By an appropriate shift of origin, U may be expanded in a power series in $x_1, x_2, \cdots x_n$,‡ uniformly and absolutely convergent in some region. In

---

* See the article on Spherical Harmonics in the *Encyclopedia Britannica.*

† Bateman, *ibid.*, p. 455.

‡ We consider only analytic solutions. For equations of the Laplace type all solutions must be analytic.

order to have an explicit expression for the number of terms of degree m in this series, we define

$$q_1(m) = 1, \quad q_n(m) = \sum_{s=0}^{m} q_{n-1}(s),$$

for $n = 2, 3, \cdots$; $m = 0, 1 \cdots$. Then there will be $q_n(m)$ terms of the power series of degree m. Let the homogeneous polynomial of degree m contained in the power series be denoted by h_m. Because the differential equation is homogeneous, h_m must itself be a solution. Operating on h_m by P gives a homogeneous polynomial g_m of degree $m - r$ for $m \geqq r$. g_m will contain $q_n(m - r)$ linearly independent terms. For h_m to be a solution, that is, for $P \cdot h_m$ to vanish identically, the coefficients of g_m must vanish, thus giving as many linear equations on the coefficients of h_m. It can be shown that these equations are linearly independent. (If $m < r$ there are no such equations, so we put $q_n(s) = 0$ if $s < 0$.)

Hence there will be $q_n(m) - q_n(m - r)$ (which we will denote by t) *linearly independent homogeneous polynomials of degree m which satisfy the equation $P \cdot U = 0$,* and *every homogeneous polynomial of degree m which is a solution of this equation may be expressed linearly in terms of such a fundamental set.** Furthermore, *any set of t linearly independent solutions, u_{mi}, which are homogeneous polynomials of degree m can be used as such a fundamental set, and*

$$\sum_{m=0}^{\infty} \sum_{i=1}^{t} c_{mi} u_{mi}$$

is a general solution in the neighborhood of the origin, the c's being arbitrary constants.

3. *Algebra associated with the general solution of a differential equation.* Consider the hypervariable

$$w = \sum_{j=1}^{n} x_j e_j$$

where the e's are numbers in a commutative and associative algebra, as yet undefined. Then for $m \geqq r$,

$$P(\partial/\partial x_1, \partial/\partial x_2, \cdots \partial/\partial x_n) \cdot w^m = [m!/(m-r)!]\, w^{m-r} P(e_1, e_2, \cdots e_n).$$

Hence w^m, *and all analytic functions* (expressible by power series) *of w, will satisfy the differential equation identically if and only if*

$$P(e_1, e_2, \cdots e_n) = 0 \tag{1}$$

This fundamental algebraic equation will be called the *auxiliary equation.*

* Bateman, *ibid.*

We are thus led to a study of the linear algebras formed by using $e_1, e_2, \cdots e_n$ as generators and having the following properties. It is understood that the algebra will be taken over the same field of numbers as the coefficients of the differential equation.

(a) Associative and commutative.

(b) $e_1, e_2, \cdots e_n$ satisfy (1).

(c) $e_1, e_2, \cdots e_n$ shall be the only generators, i. e., every number of the algebra is a polynomial or formal power series in $e_1, e_2, \cdots e_n$.

(d) There shall be no other relations between the generators independent of (a) and (b).

We now wish to find a set of *units* for the algebra. Let us first consider the special case of a first order differential equation. Then the auxiliary equation will be of the first degree, and a linear relation will hold between the numbers $e_1, \cdots e_n$. In other words these generators are not all independent and only $n - 1$ of them can be taken as units of the algebra. Suppose we pick $e_1, \cdots e_{n-1}$ as units. (Any other set of $n - 1$ linearly independent combinations of the numbers $e_1, \cdots e_n$ would evidently serve as well). Then all the products of the units $e_1, \cdots e_{n-1}$ disregarding order of multiplication, are linearly independent and can be taken as additional units of the algebra. For example, there will be $n(n-1)/2$ units of degree 2 in the generators, e_1^2, e_1e_2, etc. We thus get an infinite set of units, and at the same time the multiplication table is defined. It is seen that the units fall into groups according to their degree in the generators. Multiplication of a unit in a group corresponding to degree p in the generators by one of degree s gives a unit of degree $p + s$ in the generators.

For the general case of a differential equation of any order the situation is more complicated. The generators will themselves be linearly independent but some of their products will not be. The quantities $e_1^{\alpha_1} e_2^{\alpha_2} \cdots e_n^{\alpha_n}$, where $\sum \alpha_i = m$, are all linearly independent if $m < r$. Hence these quantities may all be taken as independent units, there being $q_n(m)$ of them for each $m < r$. For $m \geqq r$, (b) gives $q_n(m - r)$ relations for each m between these quantities. Hence in any case we can choose *for each m, t quantities of the form $e_1^{\alpha_1}e_2^{\alpha_2} \cdots e_n^{\alpha_n}$ as independent units of the algebra* (t having the same value as in Sec. 2). We denote these units by e_{mi}.

Now if we write w^m in terms of the units so defined, we get

$$w^m = \sum_{i=1}^{t} f_{mi}(x_1, x_2, \cdots x_n) e_{mi}, \qquad (m = 0, 1, \cdots).$$

It is easily shown that the coefficients of the units, f_{mi} are linearly inde-

pendent. They will also satisfy the differential equation $P \cdot U = 0$, since w^m satisfies the equation identically, by virtue of (1). Hence they may be taken as a fundamental set of solutions of degree m. Therefore

$$\sum_{m=0}^{\infty} \sum_{i=1}^{t} f_{mi} C_{mi},$$

where the C's are arbitrary constants, is a general solution of the differential equation, valid near the origin. But this is just the expansion of $(1-w)^{-1}$ with the e's replaced by arbitrary constants. There follows the theorem:

If w is defined by conditions (a), (b), (c) and (d), any solution of the equation $P \cdot U = 0$ may be expressed as $(1-w)^{-1}$ (or any other particular function of w whose expansion in powers of w contains all powers and with real coefficients) with the units replaced by suitable constants.

It is evident that the method can be extended at once to a system of equations in one dependent variable. Eq. 1 will be replaced by several equations, and the number of independent units of a particular degree in the generators will be correspondingly reduced. The method may also be extended to systems in more than one dependent variable.

As a simple example of the method in this section, we may consider the first order equation

$$\partial w/\partial x + 2\partial w/\partial y - \partial w/\partial z = 0.$$

The auxiliary equation is

$$e_1 + 2e_2 - e_3 = 0.$$

Taking e_1 and e_2 as independent units, e_3 is thus expressed in terms of them, and the hypervariable becomes

$$w = (x+z)e_1 + (y+2z)e_2.$$

Following our general theorem we consider the series $\sum w^m$. Since every product of e_1 and e_2 is linearly independent of the others, and is thus a unit, when the units are all replaced by arbitrary constants there is obtained merely a general power series in the variables $x+z$ and $y+2z$. Hence we obtain the usual general solution

$$F(x+z,\ y+2z)$$

where F is an arbitrary analytic function. It is easily seen that for every first order equation the solution is expressible in a similar form.

As a second example, consider the equation

$$\partial^3 w/\partial x \partial y \partial z = 0.$$

The auxiliary equation is $e_1 e_2 e_3 = 0$. Hence all products of order 3 or more

including all three generators are zero. The algebra is the direct sum * of three subalgebras, whose units are the various products of e_1 and e_2, e_1 and e_3, and of e_2 and e_3 respectively, all these products being independent. Forming a power series in w, we see that no terms involving x, y, and z will appear, but every term involving any two of them will be present. We are thus led to the usual solution

$$F_1(x, y) + F_2(y, z) + F_3(x, z)$$

where F_1, F_2, F_3 are arbitrary analytic functions.

Whenever the differential equation is such that the auxiliary equation has two different real factors, the algebra will be the direct sum of two corresponding subalgebras, and the general solution will be the sum of two arbitrary functions of the separate hypervariables. If a factor is repeated there will be corresponding nilpotent elements in the algebra.

4. *Special forms of the algebra associated with a differential equation.* If we exclude differential equations of the form

$$(\partial^n/\partial x_1\,\partial x_2 \cdots \partial x_n) \cdot U = 0,$$

it will be possible to replace one of the generators by the unit 1 in the auxiliary equation without reducing its degree. Suppose this can be done with e_1. We impose the same conditions (a), (b), (c), and (d) on the algebra, there being in this case, of course, only $n-1$ generators $e_2, e_3, \cdots e_n$. Hence we have the same algebra as before with n replaced by $n-1$. But w^m will now depend not only on the units for which the sum of the exponents of the generators is m, but on all the units for which the sum is less than or equal to m, thus:

$$w^m = \sum_{j=0}^{m} \sum_{i=1}^{s} f_{mji}(x_1, x_2, \cdots x_n) e_{ji}, \qquad (m = 0, 1, 2, \cdots),$$

where $s = q_{n-1}(j) - q_{n-1}(j-r)$. Therefore for each m there will be t functions f_{mji}, all linearly independent. Hence the functions f_{mji} may be taken as a fundamental set of solutions of degree m of $P \cdot U = 0$.

Therefore

$$\sum_{m=0}^{\infty} \sum_{j=0}^{m} \sum_{i=1}^{s} f_{mji} C_{mji},$$

where the C's are arbitrary constants, is a general solution of the differential equation, valid near the origin.

For example, consider the equation

$$\partial^2 u/\partial x \partial y - \partial^2 u/\partial y^2 = 0.$$

---

* Dickson, *Algebras and Their Arithmetics*, p. 33.

The auxiliary equation is

$$e_1 e_2 - e_2^2 = 0.$$

e_2 cannot be taken as the principal unit 1, because in that case the equation $e_1 - 1 = 0$ is of degree one while the auxiliary equation is of degree two. On the other hand if e_1 is replaced by 1, we get $e_2 - e_2^2 = 0$, as the equation to be satisfied by e_2. The hypervariable is

$$w = x + y e_2.$$

The algebra is finite, containing only the two units 1 and e_2. On forming w^m,

$$w^m = x^m + [(x+y)^m - x^m] e_2.$$

From the above theorem the general solution is thus obtained in the form

$$u = f(x) + g(x+y) - g(x)$$

which can, of course, be condensed into the form

$$u = F(x) + g(x+y)$$

where f, g and F are arbitrary analytic functions.

For the equations which will be treated in the succeeding sections it is possible to transform the algebra so that the square of each unit has a component in the direction of the principal unit $e_1 = 1$, but no product of two different units involves e_1. In this case, for any power series in w, if

$$\alpha_m = \sum_{j=0}^{m} \sum_{i=1}^{s} a_{mji} e_{ji} \quad \text{and if} \quad e_{ji}^2 = \gamma_{ji} e_1 + \cdots,$$

then

$$\sum_{m=0}^{\infty} \alpha_m w^m = \sum_{m=0}^{\infty} \sum_{j=0}^{m} \sum_{i=1}^{s} a_{mji} f_{mji} \gamma_{ji} e_1 + \cdots.$$

Hence *the real part* (the coefficient of the principal unit e_1), *of a power series in w is a general solution of the differential equation*, valid near the origin.

Thus far we have considered only the formal connection between the general solution of a differential equation and its algebra. The theoretical value of the ordinary complex variable z in the case of Laplace's equation in two variables lies in the theorem that any solution of the equation is the real part of some *analytic* function of z.

The proof consists in showing that to every function u which satisfies Laplace's equation there corresponds a function v such that u and v satisfy the Cauchy-Riemann equations. The equations which correspond to the Cauchy-Riemann equations for a hypervariable (that is, which express the condition for a unique differential coefficient) were first obtained by Scheffers.*

* *Leipziger Berichte*, Vol. 45 (1893), p. 828.

For an algebra with an infinity of units the Scheffers equations are also infinite in number, and it is difficult to treat such general cases. But in special cases, in particular those cases considered in the succeeding sections, it is possible to prove as in complex variables, that given any solution of a differential equation, other solutions can be determined which satisfy the Scheffers equations of the corresponding hypervariable, and thus form an analytic function of the hypervariable.

5. *Applications to particular equations.* From the original definition of t we see that it is independent of m if and only if there are two independent variables. When t is a function of m the number of coefficients f_{mjt} for w^m increases with m, and the algebra will contain an infinity of units. Thus, for two independent variables there is a finite algebra, but for more than two independent variables there is an infinity of units.

Laplace's equation in two variables is a very special equation because t is 2 and the number of units equals the number of variables. The number of units can equal the number of variables only in the case of a second order equation in two variables.

For Laplace's equation in three variables, $t = 2m + 1$. The corresponding hypervariable has already been defined above.

For Airy's fundamental equation of elasticity, namely

$$(\partial^2/\partial x^2 + \partial^2/\partial y^2)^2 \cdot U = 0,$$

t is 4. The variable

$$w = x + ye_3 \quad \text{where} \quad e_3{}^2 = 1 + 2ie_3, \quad [i = (-1)^{1/2}],$$

has units satisfying the conditions of the last section. This algebra containing the four units $e_1 = 1$, $e_2 = i$, e_3, $e_4 = ie_3$, is the simplest complex nilpotent algebra, since the square of $e_3 - i$ vanishes. Calling this nilpotent number E, the hypervariable may be written

$$w = x + iy + yE = z + yE$$

where z is an ordinary complex variable. Then

$$w^m = z^m + myz^{m-1}E,$$

and

$$\begin{aligned} f(w) &= f(z) + yf'(z)E \\ &= f(z) - yf'(z)i + yf'(z)e_3 \end{aligned}$$

and the real part of $f(w)$, that is, the real part of $f(z)$ plus the imaginary part of $yf'(z)$, is the general solution of the differential equation. The Scheffers equations are obtained by writing

$$f(w) = f_1 + f_2 i + f_3 e_3 + f_4 i e_3,$$

the f's being functions of x and y, and finding the conditions that there exist a function $f'(w)$ such that

$$df = f'(w)\,dw$$

for all values of dw. The equations are

$$\begin{aligned} \partial f_1/\partial y &= \partial f_3/\partial x, & \partial f_3/\partial y &= \partial f_1/\partial x - 2\partial f_4/\partial x, \\ \partial f_2/\partial y &= \partial f_4/\partial x, & \partial f_4/\partial y &= \partial f_2/\partial x + 2\partial f_3/\partial x. \end{aligned}$$

On elimination between these equations, it is seen that the functions f_1, f_2, f_3, f_4 satisfy Airy's equation, and conversely it is easily shown that given any solution f_1 of Airy's equation, functions f_2, f_3, and f_4 can be determined so that the Scheffers equations are satisfied. Hence *the real part, f_1, of an arbitrary analytic function of w forms a general solution of Airy's equation.*

6. *Laplace's equation in n variables.* Let

$$\Delta_n = \sum_{j=1}^{n} \frac{\partial^2}{\partial x_j^{\,2}},$$

where $x_1, x_2, \cdots x_n$ are ordinary complex variables. The hypervariable connected with the equation $\Delta_2 \cdot U = 0$ is $x_1 + ix_2$, and for $\Delta_3 \cdot U = 0$ is

$$x_1 + ix_2 \cos\alpha + ix_3 \sin\alpha,$$

where $\cos\alpha$ and $\sin\alpha$ are units. Thus to pass from the case $n = 2$ to $n = 3$ we have substituted the hypervariable $x_2 \cos\alpha + x_3 \sin\alpha$ for x_2. We shall show that this is true in general, so that the hypervariable for $\Delta_4 \cdot U = 0$ (the wave equation) is

$$w = x_1 + ix_2 \cos\alpha + ix_3 \sin\alpha \cos\beta + ix_4 \sin\alpha \sin\beta, \text{ etc.}$$

It will be shown that *any solution of $\Delta_n \cdot U = 0$ is the scalar part* (ordinary complex coefficient of the principal unit 1) *of some analytic function of the corresponding hypervariable.*

For values of n larger than 3 the Scheffers equations become too complicated to integrate directly, so the proof will be effected by a process of induction. For simplicity of notation the induction will only be carried from $n = 3$ to $n = 4$, but it will be seen that the method is general.

Let $$\omega = x_3 \cos\beta + x_4 \sin\beta.$$

We first transform ω by writing $\cos\beta$ and $\sin\beta$ as exponentials. This means only a linear transformation of the units of the algebra:

4

$$\omega = ue^{i\beta} + ve^{-i\beta} \equiv \omega',$$

where

$$2u = x_3 - ix_4, \qquad 2v = x_3 + ix_4.$$

Transforming the differential equation gives

$$\Delta_4 \cdot U = (\Delta_2 + \partial^2/\partial u\, \partial v) \cdot U \equiv \Delta'_4 \cdot U.$$

Let

$$F(\omega') = \sum_{-\infty}^{\infty} F_k(u, v) e^{ki\beta}.$$

The necessary and sufficient condition that $F(\omega')$ be analytic, is that the Scheffers equations for ω':

(2) $$\partial F_{k+1}/\partial u = \partial F_{k-1}/\partial v = F'_k,$$

where

$$F'(\omega') = \sum F'_k(u, v) e^{ki\beta},$$

should be satisfied.

Now consider any two solutions $f_{00}(x_1, x_2, u, v)$ and $f_{01}(x_1, x_2, u, v)$ of the equation $\Delta'_4 \cdot U = 0$. Calling $F_k = f_{0k}$, a set of functions f_{0k} can be determined in terms of f_{00} and f_{01} which will at the same time be solutions of $\Delta'_4 \cdot U = 0$ and also satisfy equations (2).

To prove this, consider the equations for $k = 2$:

(3) $$\partial f_{02}/\partial u = \partial f_{00}/\partial v,$$

(4) $$\partial^2 f_{02}/\partial x_1^2 + \partial^2 f_{02}/\partial x_2^2 + \partial^2 f_{02}/\partial u \partial v = 0.$$

Differentiating (3) and substituting in (4) gives

(5) $$\partial^2 f_{02}/\partial x_1^2 + \partial^2 f_{02}/\partial x_2^2 + \partial^2 f_{00}/\partial v^2 = 0.$$

(3) and (5) are equivalent to (3) and (4). Now f_{02} determined by (3) leaves $f_{02}(x_1, x_2, u_0, v)$ arbitrary, u_0 being a particular value of u. But since (5) contains no u derivatives, if it is satisfied for a particular value of u, say u_0, then it will be satisfied for all values of u. Hence if $f_{02}(x_1, x_2, u_0, v)$ is determined by means of (5), f_{02} and its x_1 derivative still being arbitrary for $x_1 = x_{10}$ and $u = u_0$, f_{02} thus found will satisfy (3) and (5) identically.

We may thus determine successively the functions f_{0k} in terms of f_{00} and f_{01}. In each case f_{0k} and its x_1 derivative are arbitrary at $x_1 = x_{10}$ and $u = u_0$ for $k > 0$, and at $x_1 = x_{10}$ and $v = v_0$ for $k < 0$.

Hence

$$f_0(x_1, x_2, \omega') \equiv \sum_{-\infty}^{\infty} f_{0k}(x_1, x_2, u, v) e^{ki\beta}$$

will be an analytic function of x_1, x_2, and ω'. Now we have

$$(6) \qquad \partial^2 f_0/\partial x_1^2 = \sum (\partial^2 f_{0k}/\partial x_1^2)\, e^{ki\beta},$$

and a similar relation (7) with respect to x_2. Since the derivative of an analytic function of ω' is itself analytic, from (2)

$$(8) \qquad \partial^2 f_0/\partial \omega'^2 = \sum (\partial^2 f_{0k}/\partial u \partial v)\, e^{ki\beta}.$$

Adding (6), (7) and (8), since f_{0k} are solutions of $\Delta'_4 \cdot U = 0$,

$$\partial^2 f_0/\partial x_1^2 + \partial^2 f_0/\partial x_2^2 + \partial^2 f_0/\partial \omega'^2 = 0.$$

The problem has thus been reduced one stage. We already know that given an analytic function $f_0(x_1, x_2, \omega')$ satisfying this equation there can be found functions f_m such that $\sum f_m e^{mi\alpha}$ is an analytic function of

$$x_1 + ix_2 \cos\alpha + i\omega' \sin\alpha.^*$$

It has thus been proved that any function $f_{00}(x_1, x_2, u, v)$ which satisfies $\Delta'_4 \cdot U = 0$ is the scalar part of some analytic function $f(w')$ of

$$w' = x_1 + ix_2 \cos\alpha + iu \sin\alpha e^{i\beta} + iv \sin\alpha e^{-i\beta}.$$

Now suppose $F_{00}(x_1, x_2, x_3, x_4)$ is any solution of $\Delta_4 \cdot U = 0$. Expressing x_3 and x_4 in terms of u and v we get a function $f_{00}(x_1, x_2, u, v)$ which satisfies $\Delta_4' \cdot U = 0$. By the above theorem this function will be the scalar part of a function $f(w')$. If $f(w')$ and w' are transformed by writing u and v in terms of x_3 and x_4, and changing the exponentials back to sines and cosines, we obtain w and $F(w)$. The scalar part remains unchanged under this transformation, so that F_{00} will be the scalar part of $F(w)$. Hence we obtain the final theorem that any solution of $\Delta_4 \cdot U = 0$ is the scalar part of some analytic function of w. This is, of course, the wave equation with $x_1 = it$. *For the ordinary form of the wave equation:*

$$\Delta_3 U - \partial^2 U/\partial t^2 = 0,$$

the hypervariable is

$$w = t + x_1 e_{10} + x_2 e_{-11} + x_3 e_{-1-1}$$

where the double infinity of units e_{ij} has a multiplication table determined by the relations

$$e_{\pm i \pm j} = \begin{matrix}\cos\\ \vdots \\ \sin\end{matrix}\; i\alpha \; \begin{matrix}\cos\\ \\ \sin\end{matrix}\; j\beta.$$

For example,

---

* This was of course proved only for the case where ω' was real or complex. But the proof depends only on the fact that the functions are analytic, so the results may be generalized at once to any hypervariable ω'.

$$e_{-1-4}e_{-21} = \sin\alpha \sin 4\beta \sin 2\alpha \cos\beta$$
$$= \tfrac{1}{4}(\cos\alpha \sin 5\beta + \cos\alpha \sin 3\beta - \cos 3\alpha \sin 5\beta - \cos 3\alpha \sin 3\beta)$$
$$= \tfrac{1}{4}(e_{1-5} + e_{1-3} - e_{3-5} - e_{3-3}).$$

Also,
$$w^2 = (t + \tfrac{1}{2}x_1^2 + \tfrac{1}{4}x_2^2 + \tfrac{1}{4}x_3^2) + 2x_1te_{10}$$
$$+ (\tfrac{1}{2}x_1^2 - \tfrac{1}{4}x_2^2 - \tfrac{1}{4}x_3^2)e_{20} + \tfrac{1}{4}(x_2^2 - x_3^2)e_{02}$$
$$+ \tfrac{1}{2}x_2x_3e_{0-2} + 2x_2te_{-11} + 2x_3te_{-1-1} - \tfrac{1}{2}x_2x_3e_{2-2}$$
$$+ x_1x_2e_{-21} + x_1x_3e_{-2-1} + \tfrac{1}{4}(-x_2^2 + x_3^2)e_{22}.$$

Each of the coefficients of the units satisfies the wave equation. These coefficients and the coefficients for the higher powers are *natural generalizations of spherical harmonics to the four dimensional space* of special relativity in which distance is given by the formula

$$s^2 = t^2 - x_1^2 - x_2^2 - x_3^2.$$

In a similar way we can generalize spherical harmonics to n dimensions.

7. *Simultaneous partial differential equations.* The two conjugate functions of the analytic function of the ordinary complex variable z constitute a general solution of the Cauchy-Riemann equations. If the real part of $f(z)$ is taken as the x component and the imaginary part as the negative of the y component of a vector field, the Cauchy-Riemann equations express the fact that the curl and divergence of the field vanishes.

In three dimensions the condition that the curl and divergence of a vector field vanish gives the following linear equations on the three components α, β and γ.*

$$\partial\gamma/\partial y = \partial\beta/\partial z, \qquad \partial\beta/\partial x = \partial\alpha/\partial y,$$
$$\partial\alpha/\partial z = \partial\gamma/\partial x, \qquad \partial\alpha/\partial x + \partial\beta/\partial y + \partial\gamma/\partial z = 0.$$

In physics these equations play a rôle intermediate between the Cauchy-Riemann and Maxwell's equations.

If w is the previously defined hypervariable belonging to Laplace's equation in three variables, and if $f(w) = \Sigma f_k e_k$ is any analytic function of w, then *a general solution of the four equations on α, β and γ is obtained by putting*

$$(9) \qquad \alpha = f_0, \qquad \beta = -f_1/2, \qquad \gamma = -f_{-1}/2.$$

That is, for any three functions α, β and γ satisfying these four equations, functions f_k can be found satisfying (9) and such that $\Sigma f_k e_k$ is an analytic function of w.

UNIVERSITY OF ILLINOIS.

* See Hanni, *Tôhoku Mathematical Journal*, Vol. 5 (1914), p. 145.

A REMAINDER FOR THE EULER-MACLAURIN SUMMATION FORMULA IN TWO INDEPENDENT VARIABLES.*

By WILLIAM DOWELL BATEN.

The object of this article is to develop a remainder for the Euler-MacLaurin summation formula in two independent variables. The first part defines and develops Bernoulli product polynomials for two independent variables. The second part gives a derivation of Euler-MacLaurin summation formula for two variables by employing certain properties of the polynomials obtained in the first part. A remainder is developed in terms of definite integrals and in terms of derivatives.

Conditions for convergence are given when the summation extends to infinity for both variables.

Krause studied this problem and developed a remainder for this summation formula in two variables for functions which can be expressed in a Taylor series. It appears that he was not satisfied with the applicability of his remainder term. The remainder given in this article is much less complicated and more practical than that obtained by Krause.† The remainder for the summation formula for two variables is not as simple as that given by Steffensen for one variable.‡

1. *Bernoulli product polynomials.* Let $B_{m,n}(x,y)$ be a polynomial in two independent variables x and y such that

$$\Delta_x\Delta_y B_{m,n}(x,y) = mnx^{m-1}y^{n-1}, \quad \text{and} \tag{1}$$

$$D_x{}^iD_y{}^jB_{m,n}(x,y) = m(i)n(j)B_{m-i,n-j}(x,y),\S \tag{2}$$

where Δ_x and Δ_y represent finite differences, $D_z{}^s$ represents the s-th derivative with respect to z, and the first subscript of B denotes the degree of the polynomial in x while the second denotes the degree of the polynomial in y.

By Taylor's theorem

---

* Presented to the Society December 31, 1930, at Cleveland.

† Krause, "Über Bernoullische Zahlen und Funktionen im Gebiete der Funktionen zweier veränderlichen Grossen," *Berichte über die Verhandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften zu Leipzig*, Vol. 55-56 (1903), pp. 39-62.

‡ Steffensen, *Interpolation*, pp. 119-138.

§ $m(r) = m(m-1)(m-2)\cdots(m-r+1) = (m!)/(m-r)!$.

$$\begin{aligned} B_{m,n}(x+h, y+k) = {} & B_{m,n}(x,y) + hD_xB_{m,n}(x,y) + kD_yB_{m,n}(x,y) \\ & + (1/2)\,(h^2D_x{}^2B_{m,n}(x,y) + 2hkD_xD_yB_{m,n}(x,y) \\ & + k^2D_y{}^2B_{m,n}(x,y)) \\ & + \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ = {} & \sum_{j=0}^{n} \sum_{i=0}^{m} (1/j!)\,(1/i!)\,h^ik^jD_x{}^iD_y{}^jB_{m,n}(x,y). \end{aligned}$$

By property (2)

$$B_{m,n}(x+h, y+k) = \sum_{j=0}^{n} \sum_{i=0}^{m} \binom{m}{i}\binom{n}{j} h^ik^jB_{m-i,n-j}(x,y).$$

By writing $m-i$ for i and $n-j$ for j the above becomes

$$B_{m,n}(x+h, y+k) = \sum_{j=0}^{n} \sum_{i=0}^{m} \binom{m}{i}\binom{n}{j} h^{m-i}k^{n-j}B_{i,j}(x,y). \tag{3}$$

By use of (3) for h and k equal to 1 and by (1)

$$\begin{aligned} & B_{m,n}(x+1, y+1) - B_{m,n}(x+1, y) - B_{m,n}(x, y+1) + B_{m,n}(x,y) \\ & = \sum_{j=0}^{n-1} \sum_{i=0}^{m-1} \binom{m}{i}\binom{n}{j} B_{i,j}(x,y) = \Delta_x\Delta_yB_{m,n}(x,y) = mnx^{m-1}y^{n-1}. \end{aligned}$$

Or

$$\sum_{j=0}^{n-1} \sum_{i=0}^{m-1} \binom{m}{i}\binom{n}{j} B_{i,j}(x,y) = mnx^{m-1}y^{n-1}. \tag{4}$$

For various values of m and n these polynomials become

$$\left\{\begin{aligned} & B_{0,0}(x,y) = 1, \\ & B_{1,0}(x,y) = x - 1/2, \qquad B_{0,1}(x,y) = y - 1/2, \\ & B_{1,1}(x,y) = xy - x/2 - y/2 + 1/4, \\ & B_{2,1}(x,y) = x^2y - xy - x^2/2 + x/2 + y/6 - 1/12, \\ & B_{1,2}(x,y) = xy^2 - xy - y^2/2 + y/2 + x/6 - 1/12, \\ & B_{2,0}(x,y) = x^2 - x + 1/6, \qquad B_{0,2}(x,y) = y^2 - y + 1/6, \\ & B_{2,2}(x,y) = x^2y^2 - xy^2 - x^2y + xy + x^2/6 + y^2/6 - x/6 - y/6 + 1/36, \\ & \text{Etc.} \end{aligned}\right. \tag{5}$$

From the above polynomials it is seen that

$$B_{i,j}(x,y) = B_{i,0}(x,y) \cdot B_{0,j}(x,y),$$

where i and j run from 0 to 2. This is evident for any i and j from (1) and (2) since x and y are independent. This shows that these polynomials for two variables x and y are the products of Bernoulli polynomials for the single variables. $B_{i,0}(x,y)$ is a Bernoulli polynomial in the single variable x while $B_{0,j}(x,y)$ is a Bernoulli polynomial in the single variable y. The polynomials in the two variables will be called Bernoulli product polynomials.

The values of the product polynomials for $x = y = 0$ will be called Bernoulli product numbers, and will be represented by the symbol

$$B_{m,n}$$

where m and n take various values. These Bernoulli product numbers may be obtained from equations 5 by substitution of zero for x and y. But they may be treated independently by use of (4) for $x = y = 0$ as

$$B_{0,0} = 1; \qquad \sum_{j=0}^{n} \sum_{i=0}^{m} \binom{m}{i} \binom{n}{j} B_{i,j} = B_{m,n}, \qquad (m, n > 1).$$

This may be written in a symbolical form as

$$\{(B_{0,0} + 1)^m - B_{m,0}\}\{(B_{0,0} + 1)^n - B_{0,n}\},$$

after the expansion the exponents of the first brackets are changed to the first subscript of the B's, while the exponents of the second brackets are changed to the second subscript of the B's.

Bernoulli product polynomials may be expressed explicitly by Bernoulli product numbers by substituting in (3), $x = y = 0$, and then setting $h = x$ and $k = y$. This becomes

$$B_{m,n}(x, y) = \sum_{j=0}^{n} \sum_{i=0}^{m} \binom{m}{i} \binom{n}{j} B_{i,j} x^{m-i} y^{n-j}.$$

This may be written symbolically as

$$B_{m,n}(x, y) = (x + B)^m (y + B)^n.$$

These Bernoulli product polynomials are not the same as the Bernoulli polynomials in two variables defined by Appell,* or the Bernoulli functions in the region of two independent variables developed by Krause,† or the double Bernoulli polynomials treated by Barnes.‡ The product numbers are not the same as the Bernoulli numbers in the region of two independent variables treated by Krause or the Bernoulli double numbers defined by Barnes. These differ from the Bernoulli polynomials in two variables mentioned by Nörlund.§

* P. Appell, "Sur les functions de Bernoulli à deux variables," *Archiv der Mathematik und Physick* (3), Vol. 4 (1902), pp. 292-3.

† Krause, "Über Bernoullische Zahlen und Funktionen im Gebiete der Funktionen zweier veränderlichen Grossen," *Berichte über die Verhandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften zu Leipzig*, Vol. 55-56 (1903), pp. 39-62.

‡ Barnes, "The Theory of the Double Gamma Functions," *Philosophical Transactions*, London, 196 A (1901), pp. 271-278.

§ Nörlund, "Mémoire sur les polynomials de Bernoulli," *Acta Mathematica*, Vol. 43, p. 121.

2. *Euler-MacLaurin Summation Formula in two Variables.* Let $\bar{B}_{m,n}(x,y)$ be a periodic function of period 1 with respect to x and y such that $\bar{B}_{m,n}(x,y)$ is identical with $B_{m,n}(x,y)/m!\,n!$ for $0 \leqq x < 1$ and $0 \leqq y < 1$. This function is determined for all values of x and y, by having to satisfy the conditions

$$\bar{B}_{m,n}(x,y) = B_{m,n}(x,y)/m!\,n! \quad \text{for } 0 \leqq x < 1 \text{ and } 0 \leqq y < 1,$$
$$\bar{B}_{m,n}(x+1, y+1) = \bar{B}_{m,n}(x,y) \quad \text{for all } x \text{ and all } y.$$

If m or n is not equal to 1, $\bar{B}_{m,n}(x,y)$ is continuous for all x and for all y. But when m or n is equal to 1 this periodic function becomes discontinuous at the points (r,s), where r and s are integers. $\bar{B}_{m,n}(x,y)$ is continuous in x if m is not equal to 1 and continuous in y if n is not equal to 1. Hence

$$D_x{}^i D_y{}^j \bar{B}_{m,n}(x,y) = \bar{B}_{m-i,n-j}(x,y), \quad \text{for} \quad m-i,\ n-j > 0.$$

Hence $\bar{B}_{m,n}(x,y)$ possesses continuous differential coefficients of x and of y of all orders up to $m-2$ for x and up to $n-2$ for y, while the differential coefficients of order $m-1$ for x and $n-1$ for y are discontinuous at the points (r,s), where r and s are integers, and also for the points where either r or s is an integer. Yet $\bar{B}_{m,1}(x,y)$ has continuous derivatives of all orders up to $m-2$ in x, while $\bar{B}_{1,n}(x,y)$ has continuous derivatives of all orders up to $n-2$ in y.

Consider the double integral

$$R_{m,n} = \int_0^1 \int_0^1 \bar{B}_{m,n}(\Theta - t, \Psi - k) f_{m,n}(x+t, y+k)\, dt\, dk,$$

where the m in the subscript of the function f represents the m-th derivative with respect to t and the n in the subscript of f means the n-th derivative of f with respect to k. f is a continuous function in x and y and possesses continuous derivatives of all orders in x and y, which are integrable.

Integrate by parts with respect to t and then with respect to k

$$R_{m,n} = \bar{B}_{m,n}(\Theta, \Psi)\Delta_x \Delta_y f_{m-1,n-1}(x,y) + P_{m,n-1} + R_{m-1,n},$$

where

$$P_{m,n-1} = \int_0^1 \bar{B}_{m,n-1}(\Theta, \Psi - k)\Delta_x f_{m-1,n-1}(x, y+k)\, dk.$$

By repeating this for $R_{m-1,n}$, $R_{m-2,n}$, etc.,

$$R_{m,n} = \sum_{v=2}^{m} \bar{B}_{v,n}(\Theta, \Psi)\Delta_x \Delta_y f_{v-1,n-1}(x,y) + \sum_{v=2}^{m} P_{v,n-1} + R_{1,n}.$$

$$P_{m,n-1} = \bar{B}_{m,n-1}(\Theta, \Psi)\Delta_x \Delta_y f_{m-1,n-2}(x,y) + P_{m,n-2}.$$

By continuing this process for the P's

$$P_{m,n-1}=\sum_{s=2}^{n-1}\bar{B}_{m,s}(\Theta,\Psi)\Delta_x\Delta_y f_{m,s-1}(x,y)+P_{m,1}.$$

$$P_{m,1}=\int_0^1\bar{B}_{m,1}(\Theta,\Psi-k)\,\Delta_x f_{m-1,1}(x,y+k)\,dk.$$

Now it is assumed that

$$0\leqq\Theta\leqq 1,\qquad 0\leqq\Psi\leqq 1.$$

From the nature of $\bar{B}_{m,n}$ and $B_{m,n}$, $\bar{B}_{m,n}=\bar{B}_{m,0}\cdot\bar{B}_{0,n}$. As $B_{0,1}(\Theta,\Psi-k)$ is discontinuous for $k=\Psi$, $P_{m,1}$ can be written in the form

$$P_{m,1}=(\bar{B}_{m,0})\{\int_0^{\Psi}\bar{B}_{0,1}(\Theta,\Psi-k)\Delta_x f_{m-1,1}(x,y+k)\,dk$$
$$+\int_{\Psi}^1\bar{B}_{0,1}(\Theta,\Psi-k+1)\Delta_x f_{m-1,1}(x,y+k)\,dk\},$$

whence, as $B_{0,1}(x,y)=y-1/2$ for $0\leqq y<1$,

$$P_{m,1}=(\bar{B}_{m,1})\{-\int_0^1 k\Delta_x f_{m-1,1}(x,y+k)\,dk$$
$$+(\Psi-1/2)\int_0^1\Delta_x f_{m-1,1}(x,y+k)\,dk+\int_{\Psi}^1\Delta_x f_{m-1,1}(x,y+k)\,dk\}$$
$$=(\bar{B}_{m,1})\{-\Delta_x f_{m-1,0}(x,y+\Psi)+\bar{B}_{0,1}(\Theta,\Psi)\Delta_x\Delta_y f_{m-1,0}(x,y)$$
$$+\int_y^{y+1}\Delta_x f_{m-1,0}(x,k)\,dk\}.$$

Therefore

$$\sum_{v=2}^{m}P_{v,n-1}=\sum_{v=2}^{m}\sum_{s=1}^{n-1}\bar{B}_{v,s}(\Theta,\Psi)\Delta_x\Delta_y f_{v-1,s-1}(x,y)$$
$$-\sum_{m}^{v=2}\bar{B}_{v,0}(\Theta,\Psi)\Delta_x f_{v-1,0}(x,y+k)+\sum_{v=2}^{m}\bar{B}_{v,0}\int_y^{y+1}\Delta_x f_{m-1,0}(x,k)\,dk.$$

$$R_{1,n}=\int_0^1\int_0^1\bar{B}_{1,n}(\Theta-t,\Psi-k)f_{1,n}(x+t,y+k)\,dt\,dk$$
$$=\int_0^1\bar{B}_{1,n}(\Theta-t,\Psi)\Delta_y f_{1,n-1}(x+t,y)\,dt+R_{1,n-1}$$
$$=\bar{B}_{0,n}(\Theta,\Psi)\int_0^1\bar{B}_{1,0}(\Theta-t,\Psi)\Delta_y f_{1,n-1}(x+t,y)\,dt+R_{1,n-1}$$
$$=(\bar{B}_{0,n}(\Theta,\Psi))\{-\Delta_y f_{0,n-1}(x+\Theta,y)+\bar{B}_{1,0}(\Theta,\Psi)\Delta_x\Delta_y f_{0,n-1}(x,y)$$
$$+\int_x^{x+1}\Delta_y f_{0,n-1}(t,y)\,dt\}+R_{1,n-1}$$

$$= -\bar{B}_{0,n}(\Theta,\Psi)\Delta_y f_{0,n-1}(x+\Theta,y)+\bar{B}_{1,n}(\Theta,\Psi)\Delta_x\Delta_y f_{0,n-1}(x,y)$$
$$+\bar{B}_{0,n}(\Theta,\Psi)\int_x^{x+1}\Delta_y f_{0,n-1}(t,y)\,dt+R_{1,n-1}.$$

Therefore

$$R_{1,n} = -\sum_{s=1}^{n}\bar{B}_{0,s}(\Theta,\Psi)\Delta_y f_{0,s-1}(x+\Theta,y)+\sum_{s=1}^{n}\bar{B}_{1,s}(\Theta,\Psi)\Delta_x\Delta_y f_{0,s-1}(x,y)$$
$$+\sum_{s=1}^{n}\bar{B}_{0,s}(\Theta,\Psi)\int_x^{x+1}\Delta_y f_{0,s-1}(t,y)\,dt+f(x+\Theta,y+\Psi)$$
$$-\bar{B}_{1,0}(\Theta,\Psi)\Delta_x f_{0,0}(x,y+\Psi)-\int_x^{x+1}f(t,y+\Psi)\,dt-\int_y^{y+1}f(x+\Theta,k)\,dk$$
$$-\bar{B}_{1,0}\int_y^{y+1}\Delta_x f(x,k)\,dk+\int_x^{x+1}\int_y^{y+1}f(t,k)\,dt\,dk,$$

after $R_{1,1}$ is treated similarly to $P_{m,1}$. Therefore

$$R_{m,n}=\sum_{v=1}^{m}\sum_{s=1}^{n}\bar{B}_{v,s}(\Theta,\Psi)\Delta_x\Delta_y f_{v-1,s-1}(x,y)-\sum_{v=1}^{m}\bar{B}_{v,0}(\Theta,\Psi)\Delta_x f_{v-1,0}(x,y+\Psi)$$
$$-\sum_{s=1}^{n}\bar{B}_{0,s}(\Theta,\Psi)\Delta_y f_{0,s-1}(x+\Theta,y)+\sum_{v=1}^{m}\bar{B}_{v,0}(\Theta,\Psi)\int_y^{y+1}\Delta_x f_{v-1,0}(x,y)\,dy$$
$$+\sum_{s=1}^{n}\bar{B}_{0,s}(\Theta,\Psi)\int_x^{x+1}\Delta_y f_{0,s-1}(t,y)\,dt-\int_x^{x+1}f(t,y+\Psi)\,dt$$
$$-\int_y^{y+1}f(x+\Theta,k)\,dk+\int_x^{x+1}\int_y^{y+1}f(t,k)\,dt\,dk+f(x+\Theta,y+\Psi).$$

Find the sum from $x=0$ to $w-1$ and then from $y=0$ to $L-1$. This becomes after transposing $f(x+\Theta,y+\Psi)$ and $R_{m,n}$

$$(6)\quad \sum_{y=0}^{L-1}\sum_{x=0}^{w-1}f(x+\Theta,y+\Psi)=-\sum_{v=1}^{m}\sum_{s=1}^{n}\bar{B}_{v,s}(\Theta,\Psi)f_{v-1,s-1}(x,y)\Big]_0^L\Big]_0^w$$
$$-\sum_{s=1}^{n}\bar{B}_{0,s}(\Theta,\Psi)\int_0^w f_{0,s-1}(t,y)\,dt\Big]_0^L$$
$$-\sum_{v=1}^{m}\bar{B}_{v,0}(\Theta,\Psi)\int_0^L f_{v-1,0}(x,y)\,dy\Big]_0^w$$
$$+\sum_{v=1}^{m}\sum_{y=0}^{L-1}\bar{B}_{v,0}(\Theta,\Psi)f_{v-1,0}(x,y+k)\Big]_0^w$$
$$+\sum_{x=0}^{w-1}\sum_{s=1}^{n}\bar{B}_{0,s}(\Theta,\Psi)f_{0,s-1}(x+\Theta,y)\Big]_0^L$$
$$+\sum_{y=0}^{L-1}\int_0^w f(t,y+\Psi)\,dt+\sum_{x=0}^{w-1}\int_0^L f(x+\Theta,k)$$
$$-\int_0^w\int_0^L f(t,k)\,dk\,dt-\sum_{y=0}^{L-1}\sum_{x=0}^{w-1}R_{m,n}.$$

Apply the Euler-MacLaurin summation formula for a single variable to the summations

$$\sum_{y=0}^{L-1} f_{v-1,0}(x, y+\Psi)\Big]_0^w, \qquad \sum_{x=0}^{w-1} f_{0,s-1}(x+\Theta, y)\Big]_0^L,$$

$$\sum_{y=0}^{L-1} \int_0^w f(t, y+\Psi)\,dt, \qquad \sum_{x=0}^{w-1} \int_0^L f(\Theta+x, k)\,dk;$$

these summations become

$$\sum_{y=0}^{L-1} \sum_{v=1}^{m} \bar{B}_{v,0}(\Theta,\Psi) f_{v-1,0}(x, y+\Psi)\Big]_0^w = \sum_{v=1}^{m} \bar{B}_{v,0}(\Theta,\Psi) \int_0^L f_{v-1,0}(x,k)\,dk\Big]_0^w$$
$$+\sum_{v=1}^{m} \sum_{s=1}^{n} \bar{B}_{v,s}(\Theta,\Psi) f_{v-1,s-1}(x,y)\Big]_0^w\Big]_0^L$$
$$-\sum_{v=1}^{m} \int_0^L \bar{B}_{v,n}(\Theta, \Psi-k) f_{v-1,n}(x,k)\,dk\Big]_0^w,$$

$$\sum_{x=0}^{w-1} \sum_{s=1}^{n} \bar{B}_{0,s}(\Theta,\Psi) f_{0,s-1}(x+\Theta, y)\Big]_0^L = \sum_{s=1}^{n} \bar{B}_{0,s}(\Theta,\Psi) \int_0^w f_{0,s-1}(t,y)\,dt\Big]_0^L$$
$$+\sum_{v=1}^{m} \sum_{s=1}^{n} \bar{B}_{v,s}(\Theta,\Psi) f_{v-1,s-1}(x,y)\Big]_0^L\Big]_0^w$$
$$-\sum_{s=1}^{n} \int_0^w \bar{B}_{m,s}(\Theta-t, \Psi) f_{m,s-1}(t,y)\,dt\Big]_0^L,$$

$$\sum_{y=0}^{L-1} \int_0^w f(t, y+k)\,dt = \int_0^L \int_0^w f(t,k)\,dt\,dk + \sum_{s=1}^{n} \bar{B}_{0,s}(\Theta,\Psi) \int_0^w f_{0,s-1}(t,y)\,dt\Big]_0^L$$
$$-\int_0^L \bar{B}_{0,n}(\Theta, \Psi-k) \int_0^w f_{0,n}(t,k)\,dt\,dk,$$

$$\sum_{x=0}^{w-1} \int_0^L f(x+\Theta, k)\,dk = \int_0^w \int_0^L f(t,k)\,dk\,dt + \sum_{v=1}^{m} \bar{B}_{v,0}(\Theta,\Psi) \int_0^L f_{v-1,0}(x,k)\,dk\Big]_0^w$$
$$-\int_0^w \bar{B}_{m,0}(\Theta-t, \Psi) \int_0^L f_{m,0}(t,k)\,dk\,dt.$$

Substituting these single summations in the double summation in (6) for $\sum_{y=0}^{L-1} \sum_{x=0}^{w-1} f(x+\Theta, y+\Psi)$, it becomes

$$\sum_{y=0}^{L-1} \sum_{x=0}^{w-1} f(x+\Theta, y+\Psi) = \int_0^L \int_0^w f(t,k)\,dt\,dk$$
$$+\sum_{v=1}^{m} \sum_{s=1}^{n} \bar{B}_{v,s}(\Theta,\Psi) f_{v-1,s-1}(x,y)\Big]_0^L\Big]_0^w$$
$$+\sum_{v=1}^{m} \bar{B}_{v,0}(\Theta,\Psi) \int_0^L f_{v-1,0}(x,k)\,dk\Big]_0^w$$

$$+\sum_{s=1}^{n}\bar{B}_{0,s}(\Theta,\Psi)\int_0^w f_{0,s-1}(t,y)\,dt\Big]_0^L$$

$$-\sum_{v=1}^{m}\int_0^L \bar{B}_{v,n}(\Theta,\Psi-k)f_{v-1,n}(x,k)\,dk\Big]_0^w$$

$$-\sum_{s=1}^{n}\int_0^w \bar{B}_{m,s}(\Theta-t,\Psi)f_{m,s-1}(t,y)\,dt\Big]_0^L$$

$$-\int_0^L \bar{B}_{0,n}(\Theta,\Psi-k)\int_0^w f_{0,n}(t,k)\,dt\,dk$$

$$-\int_0^w \bar{B}_{m,0}(\Theta-t,\Psi)\int_0^L f_{m,0}(t,k)\,dk\,dt\ -\sum_{y=0}^{L-1}\sum_{x=0}^{w-1}R_{m,n}$$

$$\sum_{y=0}^{L-1}\sum_{x=0}^{w-1}R_{m,n}=\sum_{y=0}^{L-1}\sum_{x=0}^{w-1}\int_0^1\int_0^1 \bar{B}_{m,n}(\Theta-t,\Psi-k)f_{m,n}(x+t,y+k)\,dt\,dk$$

$$=\sum_{y=0}^{L-1}\sum_{x=0}^{w-1}\int_y^{y+1}\int_x^{x+1} \bar{B}_{m,n}(\Theta-t,\Psi-k)f_{m,n}(t,k)\,dt\,dk$$

$$=\int_0^L\int_0^w \bar{B}_{m,n}(\Theta-t,\Psi-k)f_{m,n}(t,k)\,dt\,dk.$$

The remainder term is not $\sum_{y=0}^{L-1}\sum_{x=0}^{w-1}R_{m,n}$ but consists of the last five terms in the expression for $\sum_{y=0}^{L-1}\sum_{x=0}^{w-1}f(x+\Theta,y+\Psi)$.

If the function f, for which the sum is desired, is a polynomial in x and y the remainder terms are zero for large enough values of m and n.

Let Θ and Ψ both be equal to zero, then

$$\sum_{y=0}^{L-1}\sum_{x=0}^{w-1}f(x,y)=\int_0^L\int_0^w f(t,k)\,dt\,dk+\sum_{s=1}^{n}\sum_{v=1}^{m}\bar{B}_{v,s}f_{v-1,s-1}(x,y)\Big]_0^w\Big]_0^L$$

$$+\sum_{v=1}^{m}\bar{B}_{v,0}\int_0^L f_{v-1,0}(x,k)\,dk\Big]_0^w+\sum_{s=1}^{n}\bar{B}_{0,s}\int_0^w f_{0,s-1}(t,y)\,dt\Big]_0^L$$

$$-\sum_{v=1}^{m}\int_0^L \bar{B}_{v,n}(0,k)f_{v-1,n}(x,k)\,dk\Big]_0^w$$

$$-\sum_{s=1}^{n}\int_0^w \bar{B}_{m,s}(t,0)f_{m,s-1}(t,y)\,dt\Big]_0^L$$

$$-\int_0^L \bar{B}_{0,n}(0,k)\int_0^w f_{0,n}(t,k)\,dt\,dk$$

$$-\int_0^w \bar{B}_{m,0}(t,0)\int_0^L f_{m,0}(t,k)\,dk\,dt$$

$$-\int_0^L\int_0^w \bar{B}_{m,n}(t,k)f_{m,n}(t,k)\,dt\,dk.$$

This is the Euler-MacLaurin Summation Formula for two variables.

By using $\Delta_{\cdot t} B_{2r,0}(0,0) = B_{2r,0}(t,0) + B_{2r}$, and

$$\Delta_{\cdot t}\Delta_{\cdot k} B_{2r,2s}(0,0) = B_{2r,2s}(t,k) - B_{2r,2s}(0,k) - B_{2r,2s}(t,0) + B_{2r,2s},$$

the above becomes

$$\begin{aligned}
\sum_{y=0}^{L-1}\sum_{x=0}^{w-1} f(x,y) &= \int_0^L\int_0^w f(t,k)\,dt\,dk + \sum_{u=1}^{s-1}\sum_{v=1}^{r-1} \bar{B}_{2v,2u} f_{2v-1,2u-1}(x,y)\Big]_0^w\Big]_0^L \\
&+ \sum_{u=1}^{s-1} \bar{B}_{1,2u} f_{0,2u-1}(x,y)\Big]_0^w\Big]_0^L + \sum_{v=1}^{r-1} \bar{B}_{2v,1} f_{2v-1,0}(x,y)\Big]_0^L\Big]_0^w \\
&+ \bar{B}_{1,1} f_{0,0}(x,y)\Big]_0^w\Big]_0^L + \sum_{v=1}^{r-1} \bar{B}_{2v,0}\int_0^L f_{2v-1,0}\,(x,k)\,dk\Big]_0^w \\
&+ \bar{B}_{1,0}\int_0^L f_{0,0}(x,k)\,dk\Big]_0^w + \sum_{u=1}^{s-1} \bar{B}_{0,2u}\int_0^w f_{0,2u-1}(t,y)\,dt\Big]_0^L \\
&+ \bar{B}_{0,1}\int_0^w f_{0,0}(t,y)\,dt\Big]_0^L - \sum_{v=1}^{r-1}\int_0^L \bar{B}_{2v,0}\Delta_{\cdot k}\bar{B}_{0,2s} f_{2v-1,2s}(x,k)\,dk\Big]_0^w \\
&- \sum_{u=1}^{s-1}\int_0^w \bar{B}_{0,2u}\Delta_{\cdot t}\bar{B}_{2r,0} f_{2r,2u-1}(t,y)\,dt - \int_0^L \bar{B}_{1,0}\Delta_{\cdot k}\bar{B}_{0,2s} f_{0,2s}\,(x,k)\,dk\Big]_0^w \\
&- \int_0^w \bar{B}_{0,1}\Delta_{\cdot t}\bar{B}_{2r,0} f_{2r,0}(t,y)\,dt\Big]_0^L - \int_0^L \Delta_{\cdot k}\bar{B}_{0,2s}\Big[\int_0^w f_{0,2s}\,(t,k)\,dt\Big]dk \\
&- \int_0^w \Delta_{\cdot t}\bar{B}_{2r,0}\Big[\int_0^L f_{2r,0}(t,k)\,dk\Big]dt + \int_0^L\int_0^w \bar{B}_{2r,2s} f_{2r,2s}(t,k)\,dt\,dk.
\end{aligned}$$

By using the first theorem of the mean since $\Delta_{\cdot t}\bar{B}_{2r,0}$ is always positive, or does not change its sign,

$$\begin{aligned}
\int_0^w \bar{B}_{0,1}\Delta_{\cdot t}\bar{B}_{2r,0} f_{2r,0}(t,y)\,dt\Big]_0^L &= f_{2r,0}(g,y)\int_0^w \bar{B}_{0,1}\Delta_{\cdot t}\bar{B}_{2r,0}(0,0)\,dt\Big]_0^L \\
&= w\cdot f_{2r,0}(g,y)\int_0^1 \bar{B}_{0,1}\Delta_{\cdot t}\bar{B}_{2r,0}(0,0)\,dt \\
&= w\cdot \bar{B}_{2r,1} f_{2r,0}(g,y), \qquad (0 \leqq g \leqq w),
\end{aligned}$$

since $B_{2r,0}(t,0)$ is symmetrical to $(1/2, 0)$ and hence

$$\int_0^1 B_{2r,0}(t,y)\,dt\Big]_0^L = 0.$$

Using this idea for the integrals involving the Δ's, the above becomes

$$\sum_{y=0}^{L-1}\sum_{x=0}^{w-1} f(x,y) = \int_0^L\int_0^w f(t,k)\,dt\,dk + \sum_{u=1}^{s-1}\sum_{v=1}^{r-1} \bar{B}_{2v,2u} f_{2v-1,2u-1}(x,y)\Big]_0^w\Big]_0^L$$

$$+\sum_{v=1}^{r-1}\bar{B}_{2v,0}\int_0^L f_{2v-1,0}(x,k)\,dk\Big]_0^w+\bar{B}_{1,0}\int_0^L f(x,k)\,dk\Big]_0^w$$

$$+\sum_{u=1}^{s-1}\bar{B}_{0,2u}\int_0^w f_{0,2u-1}(t,y)\,dt\Big]_0^L+\sum_{u=1}^{s-1}\bar{B}_{1,2u}f_{0,2u-1}(x,y)\Big]_0^w\Big]_0^L$$

$$+\sum_{v=1}^{r-1}\bar{B}_{2v,1}f_{2v-1,0}(x,y)\Big]_0^L\Big]_0^w+\bar{B}_{1,1}f_{0,0}(x,y)\Big]_0^w\Big]_0^L$$

$$+\bar{B}_{0,1}\int_0^w f_{0,0}(t,y)\,dt\Big]_0^L-\sum_{v=1}^{r-1}\bar{B}_{2v,2s}L\cdot f_{2v-1,2s}(x,q_v)\Big]_0^w$$

$$-\sum_{u=1}^{s-1}\bar{B}_{2r,2u}w\cdot f_{2r,2u-1}(p_u,y)\Big]_0^L-\bar{B}_{1,2s}L\cdot f_{0,2s}(x,\bar{p}_1)\Big]_0^w$$

$$-\bar{B}_{2r,1}w\cdot f_{2r,0}(\bar{q}_2,y)\Big]_0^L-\bar{B}_{0,2s}L\cdot f_{0,2s}(\bar{q}_3,\bar{p}_3)$$

$$-\bar{B}_{2r,0}w\cdot f_{2r,0}(\bar{q}_4,\bar{p}_4)+L\cdot w\cdot\bar{B}_{2r,2s}f(\bar{q}_5,\bar{p}_6),$$

where

$$0\leqq q_i\leqq L,\qquad 0\leqq p_i\leqq w,$$
$$0\leqq\bar{q}_i\leqq L,\qquad 0\leqq\bar{p}_i\leqq w.$$

Let
$$F_{i,j}(x,y)=f_{i-1,j-1}(x,y),$$

where the first subscript represents differentiation with respect to x and the second represents differentiation with respect to y. By using this relation, equation 6 becomes

$$\sum_{y=0}^{L-1}\sum_{x=0}^{w-1}f(x,y)=\sum_{s=1}^{n}\sum_{v=1}^{m}\bar{B}_{v,s}F_{v,s}(x,y)\Big]_0^L\Big]_0^w-\sum_{v=0}^{m}\int_0^L\bar{B}_{v,n}(0,k)F_{v,n+1}(x,k)\,dk\Big]_0^w$$
$$-\int_0^w\int_0^L\bar{B}_{m,n}\,(t,0)F_{m+1,s}(t,y)\,dt\Big]\ ,$$
$$-\sum_{s=0}^{n}\int_0^w\bar{B}_{m,s}(t,0)F_{m+1,s}(t,y)\,dt\Big]_0^L\ ,$$

which makes the writing of the formula simplier.

If L and w approach infinity, sufficient conditions for the validity of this summation formula can be given. If $\sum_{y=0}^{\infty}\sum_{x=0}^{\infty}f(x+\Theta,y+\Psi)$ and $\int_0^\infty\int_0^\infty f(t,k)dt\,dk$ converge and if $f_{v-1,s-1}(w,L)$ approach zero as w and L approach infinity $(v=1,2,\cdots,m)$, $(s=1,2,\cdots,n)$, then

$$\sum_{y=0}^{\infty}\sum_{x=0}^{\infty}f(x+\Theta,y+\Psi)=\int_0^\infty\int_0^\infty f(t,k)\,dt\,dk+\sum_{s=1}^{n}\sum_{v=1}^{m}\bar{B}_{v,s}f_{v-1,s-1}(0,0)$$
$$+\sum_{v=1}^{m}\bar{B}_{v,0}\int_0^\infty f_{v-1,0}(0,k)\,dk+\sum_{s=1}^{n}\bar{B}_{0,s}\int_0^\infty f_{0,s-1}(t,0)\,d$$

$$-\sum_{v=1}^{m}\int_0^\infty \bar{B}_{v,n}(0,k)f_{v-1,n}(0,k)\,dk$$
$$-\sum_{s=1}^{n}\int_0^\infty \bar{B}_{m,s}(t,0)f_{m,s-1}(t,0)\,dt$$
$$-\int_0^\infty \bar{B}_{0,n}(0,k)\int_0^\infty f_{0,n}(t,k)\,dt\,dk$$
$$-\int_0^\infty \bar{B}_{m,0}(t,0)\int_0^\infty f_{m,0}(t,k)\,dk\,dt$$
$$-\int_0^\infty\int_0^\infty \bar{B}_{m,n}(t,k)f_{m,n}(t,k)\,dt\,dk,$$

provided $\int_0^\infty f_{0,n}(t,k)\,dk$ and $\int_0^\infty f_{m,0}(t,k)\,dt$ exist.

UNIVERSITY OF MICHIGAN.

THE CELESTIAL SPHERE.

By F. MORLEY.

1. *The Celestial Sphere.*

We are witnessing the arithmetizing of the physical universe as it now appears. The process is, on a grand scale, similar to the arithmetizing of space as it then appeared by Descartes and others.

The merger of Descartes for a plane is effected when we prove that there is a one-to-one correspondence between numbers x and the points of the plane. The numbers are of course the unrestricted numbers of algebra, usually called complex.

By an inversion from an outside point, we have the numbers x attached to the points of a sphere Ω, which we regard as lying in a euclidean space.

In terms of radius r, longitude θ, and colatitude ϕ, we may take

$$x = 2r \cot(\phi/2) \exp i\theta.$$

We now consider the interior of Ω only. The arc of a circle orthogonal to Ω is named by the end-points x and y on Ω, or collectively by a quadratic q whose zeros are x and y. The handling of arcs is then the handling of quadratics, $q_1, q_2, \cdots$.

2. *The Theory of Quadratics.*

Two ordered quadratics have a Jacobian j_{12} and a bilinear invariant q_{12}. They have then an absolute covariant j_{12}/q_{12}. The vanishing of q_{12} says that the two pairs of zeros are harmonic, whence the arcs intersect at right angles, or are ***normal***. The Jacobian j_{12} is then the common normal of q_1 and q_2.

For 4 points of Ω we have three pairs of arcs, and three common normals. The theory of the quartic shows that these three quadratics q_i are themselves normal and meet at a point at right angles.

For any three quadratics we have as in Salmon's Higher Algebra the relation or *syzygy*

$$\begin{vmatrix} q_1 & q_2 & q_3 & 0 \\ q_{11} & q_{12} & q_{13} & q_1 \\ q_{21} & q_{22} & q_{23} & q_2 \\ q_{31} & q_{32} & q_{33} & q_3 \end{vmatrix} = 0.$$

This is familiar (for real numbers) as the equation of a conic, say

$$ax^2 + by^2 + cz^2 + 2xyz + 2gzx + 2hxy = 0$$

or

$$(a/fgh - 1/f^2)x^2 + \cdot \cdot \cdot + (x/f + y/g + z/h)^2 = 0.$$

That is, quadratics fall into sets of four, three determining a fourth, such that

$$\lambda_1 q_1^2 + \lambda_2 q_2^2 + \lambda_3 q_3^2 + \lambda_4 q_4^2 = 0,$$

for any value of the variable. Now $\lambda_1 q_1^2 + \lambda_2 q_2^2 = 0$ is four points giving a pair of arcs normal to the common normal of q_1 and q_2, j_{12}. But then also $\lambda_3 q_3^2 + \lambda_4 q_4^2 = 0$ is the same four points giving a different pair of arcs normal to j_{34}. By the theory of four points the arcs j_{12} and j_{34} are normal.

The set of four arcs is then such that the common normals j_{12}, j_{34} are normal, j_{23}, j_{14} are normal, and j_{31}, j_{24} are normal. This configuration of 10 arcs, each normal to three others, is in the restricted domain of real numbers the Descartes configuration.

There is between any seven quadratics

$$q_i \equiv \alpha_i x^2 + \beta_i x + \gamma_i$$

a quadric relation

$$| \alpha_i^2 \quad \beta_i^2 \quad \gamma_i^2 \quad \beta_i\gamma_i \quad \gamma_i\alpha_i \quad \alpha_i\beta_i \quad q_i^2 | = 0,$$

and we ask the meaning of the 6-rowed condition on six quadratics

$$| \alpha_i^2 \quad \beta_i^2 \quad \gamma_i^2 \quad \beta_i\gamma_i \quad \gamma_i\alpha_i \quad \alpha_i\beta_i | = 0.$$

In terms of the zeros it is the symmetric two-to-two or biquadratic correspondence, or involution. In the restricted domain it is familiar as Pascal's theorem when the six points α_i, β_i, γ_i are on a given conic, and the theory built on this theorem by Steiner, Plücker, Kirkman, Cayley, Salmon, Veronese, Cremona and Richmond is immediately available. When the six points are not given as on a conic (that is when each q_i is not a square) the theory in the restricted domain is that of Poncelet polygons.

In the present theory Pascal's theorem takes the form:

Let there be 6 arcs q_i in a biquadratic involution,

$$(\alpha x)^2 (\alpha y)^2 + \Lambda (x - y)^2.$$

Ordering them say as 1 2 3 4 5 6, we have three pairs of common normals j_{12}, j_{45}. These three pairs have three common normals. These three have themselves a common normal. Cp. Klein, *Mathematische Annalen,* Vol. 22; *Works,* Vol. 1, p. 406.

If this happens for one ordering, it holds for all the 60 orderings. We have the 60 Pascal arcs.

It may be mentioned that the Kirkman arc for the above ordering is

$$j_{12}/q_{12} + j_{23}/q_{23} + j_{34}/q_{34} + j_{45}/q_{45} + j_{56}/q_{56} + j_{61}/q_{61}.$$

There is a cubic relation between any 11 quadratics namely

$$| \alpha_i^3 \cdots \alpha_i^2\beta_i \cdots \alpha_i\beta_i\gamma_i \cdots q_i^3 | = 0.$$

If q_{11} is arbitrary then we have 10 quadratics in a three-to-three or bicubic involution. And if also q_{10} is arbitrary, we have the fact that eight arcs determine a ninth, the relation being a mutual one.

3. *The Euclidean Case.*

There is now the interesting case when the universe Ω expands. We make the radius of the sphere Ω infinite. The arcs become lines of a euclidean space. The somewhat difficult subject of lines in a euclidean space, (see for instance Study, *Geometrie der Dynamen* and Blaschke, *Differential Geometrie*, Vol. 1, p. 264), so viewed is merely the humane theory of quadratics (under homographies). For example the configuration of 10 arcs becomes the Petersen-Morley configuration.

ON THE CONFIGURATION OF THE NINE BASE POINTS OF A PENCIL OF EQUIANHARMONIC CUBICS.

By J. YERUSHALMY.*

1. In a paper published in this Journal,† we proposed to show that every pencil of equianharmonic cubics is contained in a net of such cubics through 6 of the 9 base-points of the pencil, and that these 6 points form the vertices of two in-circumscribed triangles in a cubic ϕ which are three-fold perspective from the vertices of a third in-circumscribed triangle in the same cubic. However, the proof given there is incomplete. The purpose of this note is to complete the proof and to show that the 9 base points of a pencil of equianharmonic cubics divide into 3 triples i, j, k such that through any two triples i, j there passes a cubic ϕ_k in which they form the vertices of two in-circumscribed triangles three-fold perspective from the vertices of a third in-circumscribed triangle in ϕ_k. Thus through any two triples there passes a net of equianharmonic cubics.

2. The reasoning used in P was the following. The plane α of the pencil was mapped into a double plane π by means of a net of cubics on 7 of the 9 base-points. The branch-curve in π is a quartic, the equation of which was proved by Chisini ‡ to be

$$f_4 \equiv (\partial^2\phi_3/\partial x_3{}^2)\phi_3 - \tfrac{1}{2}(\partial\phi_3/\partial x_3)^2 = 0.$$

We considered the cubic surface

$$F_3 \equiv (\partial^2\phi_3/\partial x_3{}^2)x_4{}^2 + 2^{1/2}(\partial\phi_3/\partial x_3)x_4 + \phi_3 = 0,$$

which, when projected from the point $O \equiv (0,0,0,1)$ on it upon the plane π $(x_4 = 0)$, gives $f_4 = 0$ as branch curve. The $(1-2)$ correspondence A between α and π and the $(1-2)$ correspondence B between F_3 and π defines a $(1-1)$ correspondence D between F_3 and α which sends the sections of F_3 by the planes on O into the net of cubics on the original 7 base-points. We showed that F_3 contains a net of equinaharmonic plane sections and that

* National Research Fellow.

† "Construction of Pencils of Equianharmonic Cubics," *American Journal of Mathematics*, Vol. 53 (1931), pp. 319-332. This paper will be referred to in the sequel by the letter P.

‡ Chisini, "Sui fasci di cubiche a modulo costante," *Rendiconti del Circolo Matematico di Palermo*, Vol. 41 (1916).

it admits a cyclic homology Γ_3 of period 3 into itself. By a mapping C of F_3 into α the homology Γ_3 has as image a biquadratic transformation T cyclic of period 3 having the fundamental points of the direct and inverse transformation coincident, and from this follows the conclusion.

The weak point in the proof is that we assumed that the $(1-1)$-correspondence D between F_3 and α is necessarily a Clebsch mapping C of F_3 on α by means of which the plane sections of F_3 go into cubics on 6 of the 7 base-points of the original net. This would be so if and only if the double-tangent

$$d \equiv \partial^2\phi_3/\partial x_3^2 = 0,$$

which is the image of the point O on F_3 in the mapping B, would correspond in the mapping A to one of the 7 base-points in α. If, however, this double-tangent corresponds to a cubic degenerating into a conic on five of the base-points and a line on the other two base-points, then the correspondence D is such that either the point O goes by it into the line and the cubic f_3 cut out on F_3 by the tangent plane at O goes into the conic or vice versa. It is easily verified that in the first case D is a mapping M of F_3 unto α by means of which the plane sections of F_3 go into quartics on the 7 base-points

$$(1^2\, 2^2\, 3^1\, 4^1\, 5^1\, 6^1\, 7^1)_4$$

and in the second case D is a mapping N sending the plane sections of F_3 into quintics on the 7 base-points

$$(1^1\, 2^1\, 3^2\, 4^2\, 5^2\, 6^2\, 7^2)_5.$$

In both these mappings the net of cubics cut out on F_3 by the planes on O go into the original net of cubics on the 7 base-points $(1\,2\,3\,4\,5\,6\,7)_3$.

There remains hence the possibility of the existence of three distinct types (c), (m), (n) of pencils of equianharmonic cubics corresponding to the three different mappings. We prove, however, that there is only one type of pencil of equianharmonic cubic, by showing that the pencils of type (c), i. e. those considered in P give rise to all the three mappings, and every pencil of type (m) or (n) is also of type (c).

3. Take a pencil of equianharmonic cubics of type (c). The existence of such pencils was proved in P also independently of the cubic surface F_3 (*loc. cit.*, p. 331). Let $1, 2, \cdots, 9$ be the base-points of the pencil where $(1, 2, 3)$ $(4, 5, 6)$ are the vertices of two 3-fold perspective in-circumscribed triangles in a cubic ϕ_3, and are therefore the base-points of a periodic biquadratic transformation T

$$(1^2\, 2^2\, 3^2\, 4^1\, 5^1\, 6^1),$$

and 7, 8, 9 form a cycle for T. Consider an equianharmonic cubic ψ of the pencil. T leaves ψ invariant. Let the birational transformation of the cubic ψ into itself affected by T be

$$\gamma : u' = \epsilon u + b \qquad (u \text{ is the abelian parameter on } \psi, \; b \text{ a constant, } \epsilon^3 = 1).$$

From P, p. 329, we obtain

$$(1) \quad \begin{array}{llll} \text{a. } \epsilon u_4 + b \equiv -(u_4 + u_5); & -\epsilon(u_1 + u_3) + b \equiv u_1; & \epsilon u_7 + b \equiv u_8 \\ \text{b. } \epsilon u_5 + b \equiv -(u_5 + u_6); & -\epsilon(u_1 + u_2) + b \equiv u_2; & \epsilon u_8 + b \equiv u_9 \\ \text{c. } \epsilon u_6 + b \equiv -(u_6 + u_4); & -\epsilon(u_2 + u_3) + b \equiv u_3; & \epsilon u_9 + b \equiv u_7, \end{array}$$

or

$$(2) \quad \begin{array}{l} \text{a. } \epsilon u_4 + b = -(u_4 + u_5) + \omega_1, \\ \text{b. } \epsilon u_5 + b = -(u_5 + u_6) + \omega_2, \\ \text{c. } \epsilon u_6 + b = -(u_6 + u_4) + \omega_3, \end{array} \qquad (\omega_i \text{ are linear combinations of the periods.})$$

from which $\omega_3 = -\epsilon^2\omega_2 - \epsilon\omega_1$. Substituting for ω_3 in 2c. and adding a, b, c, we obtain

$$u_4 + u_5 + u_6 = (\epsilon - 1)b + [(1 - \epsilon^2)\omega_2 - \epsilon(1 - \epsilon^2)\omega_1]/(1 - \epsilon^2),$$

or

$$(3) \qquad u_4 + u_5 + u_6 \equiv (\epsilon - 1)b.$$

Similarly it can be shown that

$$(4) \quad \begin{array}{ll} \text{(a) } u_1 + u_2 + u_3 \equiv \epsilon(\epsilon - 1)b & \text{(b) } u_7 + u_8 + u_9 \equiv (\epsilon + 2)b \\ \text{(c) } u_5 \equiv \epsilon^2 u_4 - b & \text{(d) } u_6 \equiv \epsilon u_4 + \epsilon b \\ \text{(e) } u_2 \equiv \epsilon^2 u_1 - \epsilon b & \text{(f) } u_3 \equiv \epsilon u_1 + \epsilon^2 b. \end{array}$$

Consider now a biquadratic transformation S with fundamental points at $(1^1\,2^1\,3^1\,7^2\,8^2\,9^2)$. The fundamental points of the inverse transformation S^{-1} will generally be some other 6 points $(1_0{}^2\,2_0{}^2\,3_0{}^2\,7_0{}^1\,8_0{}^1\,9_0{}^1)$. S sends the cubics on 1, 2, 3, 7, 8, 9 into the cubics on $1_0, 2_0, 3_0, 7_0, 8_0, 9_0$. It sends ψ into some cubic ψ_0.

The homoloidal net Σ' on 1, 2, 3, 7, 8, 9 cuts out on ψ a $g_3{}^2$. But since by 4a and 4b

$$u_1 + u_2 + u_3 \equiv \epsilon(\epsilon - 1)b,$$
$$2(u_7 + u_8 + u_9) \equiv 2(\epsilon + 2)b,$$

and hence

$$u_1 + u_2 + u_3 + 2(u_7 + u_8 + u_9) \equiv 3b,$$

the sum of the abelian parameter at a set of the $g_3{}^2$ is equivalent to $-3b$. Also if we transform any set of 3 points a, b, c on a line by a transformation $\bar{\gamma} : u' = \epsilon^2 u - b$ we obtain

$$u'_a = \epsilon^2 u_a - b;\quad u'_b = \epsilon^2 u_b - b;\quad u'_c = \epsilon^2 u_c - b,$$

from which

$$u'_a + u'_b + u'_c \equiv \epsilon^2 (u_a + u_b + u_c) - 3b.$$

But since

$$u_a + u_b + u_c \equiv 0$$

$$\therefore \quad u'_a + u'_b + u'_c \equiv -3b.$$

Hence $\bar{\gamma}$ transforms the $g_3{}^2$ cut out on ψ by the lines of the plane into the $g_3{}^2$ cut out on it by the quartics of Σ'. Therefore, the product $\bar{\gamma}S$ is a birational transformation of ψ into ψ_0 sending the $g_3{}^2$ cut out on ψ by the lines of the plane into the $g_3{}^2$ cut out on ψ_0 by the lines of the plane. Hence $\bar{\gamma}S$ is involved by a collineation λ between ψ and ψ_0. The collineation sends the points $1, 2, 3, 7, 8, 9$ into the points $1_0\, 2_0\, 3_0\, 7_0\, 8_0\, 9_0$. To prove this take for example the point 1. $\bar{\gamma}$ sends it into the point $\epsilon^2 u_1 - b$, which by 4b and 4f, is the sixth point of intersection of ψ with the conic $(1, 3, 7, 8, 9)$ and this point goes by S into one of the points $1_0, 2_0, 3_0$, say 1_0. In the same way it is easily seen that λ sends 2 into 2_0 etc. Multiplying λ by S^{-1} we obtain the birational transformation $\bar{\gamma}$ which is evidently a biquadratic transformation sending ψ into itself and having the fundamental points of the direct and inverse transformation coincident which is sufficient to prove the existence of a net of equianharmonic cubics on $1, 2, 3, 7, 8, 9$. A similar reasoning shows that there exists a net of equianharmonic cubics also on $4, 5, 6, 7, 8, 9$. There are therefore defined two other cubics ϕ_2 and ϕ_1 such that each one of them has two triples as vertices of two in-circumscribed triangles 3-fold perspective from the vertices of a third in-circumscribed triangle in it. Each of the three cubics ϕ_i passes through the cusps of the 6 cuspidal cubics of the pencil (Pp. 332).

There obviously can be no more than 3 nets of equianharmonic cubics on these points. For if there were a fourth it would define another cubic ϕ which would have with one of the 3 cubics ϕ_i more than 9 points in common since it would also have to pass through the cusps and have two inscribed triangles with vertices among the 9 base-points.

4. The division of the 9 base-points into 3 triples shows that a pencil of type (c) gives rise also to the mappings M and N. For we can select 7 of the 9 base-points in 2 essentially distinct ways: either they contain 2 triples and one point of the third triple, or they consist of one riple and of two points of each of the remaining 2 triples. If we map α on π by means of cubics on 7 points containing 2 triples, then the double tangent d of $f_4 = 0$ corresponding to O on F_3 will be the image of the seventh base-points and the

correspondence D between F_3 and α is a Clebsch mapping C. If, however, the seven points contain only 1 triple, then the double tangent d is necessarily the image of a cubic composed of a conic on 5 of the 7 points and a line on the other 2 points. Then, if in the mapping of F_3 on α, we let O correspond to the line and the cubic f_3 — to the conic, we obtain the mapping M, and if we let O correspond to the conic and f_3 to the line, we obtain the mapping N.

5. To complete the proof given in P, we must show now that every pencil of type (m) is also of type (c). Take such a pencil. Let $1, 2, \cdots, 9$ be its base-points. In the mapping M between F_3 and α the plane-sections of F_3 go into the quartics

$$(1^2\, 2^2\, 3^1\, 4^1\, 5^1\, 6^1\, 7^1)_4$$

and the line (12) is the image of the point O. It can be shown that the homology Γ_3 of F_3 is mapped by M into a Cremona transformation T_7 of order 7 with 8 fundamental points, 7 of which coincide for T_7 and T_7^{-1}. The homoloidal nets are the following:

$$\begin{aligned} \Sigma' &: (1^3\, 2^4\, 3^1\, 4^2\, 5^2\, 6^2\, 7^3\, 9^1)_7, \\ \Sigma &: (1^4\, 2^3\, 3^3\, 4^2\, 5^2\, 6^2\, 7^1\, 8^1)_7. \end{aligned}$$

Knowing the configuration of the 27 lines on F_3 (they distribute in 9 triples, each triple is in a plane and the three lines concur at a point on the cubic ϕ), we can deduce the fundamental curves of T_7:

$$(5) \qquad \begin{array}{llcccl} \text{(a)} & 1 \rightarrow & (1^2\, 2^1\, 3^1\, 4^1\, 5^1\, 6^1\, 8^1)_3 & \rightarrow & (1^2\, 2^2\, 3^1\, 4^1\, 5^1\, 6^1\, 7^2\, 9^1)_4 & \rightarrow 1 \\ \text{(b)} & 2 \rightarrow & (1^2\, 2^2\, 3^2\, 4^1\, 5^1\, 6^1\, 7^1\, 8^1)_4 & \rightarrow & (1^1\, 2^2\, 4^1\, 5^1\, 6^1\, 7^1\, 9^1)_3 & \rightarrow 2 \\ \text{(c)} & 3 \rightarrow & (1\,3)_1 & \rightarrow & (1^1\, 2^2\, 3^1\, 4^1\, 5^1\, 6^1\, 7^1)_3 & \rightarrow 3 \\ \text{(d)} & 4 \rightarrow & (1\,2\,4\,3\,5)_2 & \rightarrow & (1\,2\,4\,6\,7)_2 & \rightarrow 4 \\ \text{(e)} & 5 \rightarrow & (1\,2\,5\,3\,6)_2 & \rightarrow & (1\,2\,5\,4\,7)_2 & \rightarrow 5 \\ \text{(f)} & 6 \rightarrow & (1\,2\,6\,3\,4)_2 & \rightarrow & (1\,2\,6\,5\,7)_2 & \rightarrow 6 \\ \text{(g)} & 7 \rightarrow & (1^2\, 2^1\, 3^1\, 4^1\, 5^1\, 6^1\, 7^1)_3 & \rightarrow & (2\,7)_1 & \rightarrow 7 \\ \text{(h)} & 8 \rightarrow & 9 & \rightarrow & (1\,2)_1 & \rightarrow 8 \\ \text{(j)} & 9 \rightarrow & (1\,2)_1 & \rightarrow & 8 & \rightarrow 9. \end{array}$$

The transformation T_7 leaves invariant each cubic of the original pencil. Take a cubic ψ of this pencil and let the birational transformation of ψ into itself involved by T_7 be

$$\gamma : u' = \epsilon u + b.$$

From 5c, g, h, j, we obtain

$$(6)\qquad \begin{array}{ll} \epsilon u_3 + b \equiv -(u_1+u_8), & -\epsilon(u_2+u_7)+b \equiv u_7, \\ \epsilon u_9 + b \equiv -(u_1+u_2), & -\epsilon(u_1+u_2)+b \equiv u_8, \end{array}$$

from which

$$(7)\qquad \begin{array}{ll} \text{a. } u_3 \equiv \epsilon u_1 + \epsilon b & \text{b. } u_7 \equiv \epsilon^2 u_2 - \epsilon b \\ \text{c. } u_8 \equiv -\epsilon u_1 - \epsilon u_2 + b & \text{d. } u_9 \equiv -\epsilon^2 u_1 - \epsilon^2 u_2 - \epsilon^2 b. \end{array}$$

Consider now a biquadratic transformation T_4 with fundamental points at $(1^1\, 7^1\, 8^1\, 2^2\, 3^2\, 9^2)$. The fundamental points of T_4^{-1} will be

$$(1_0{}^2\, 7_0{}^2\, 8_0{}^2\, 2_0{}^1\, 3_0{}^1\, 9_0{}^1).$$

The transformation T_4 sends ψ into ψ_0. The quartics Φ' : $(1^1\, 7^1\, 8^1\, 2^2\, 3^2\, 9^2)_4$ cut out on ψ a $g_3{}^2$. But since by (7)

$$\begin{gathered} u_1 + u_7 + u_8 \equiv (1-\epsilon)u_1 + \epsilon(\epsilon-1)u_2 + (1-\epsilon)b, \\ 2(u_2+u_3+u_9) \equiv 2\epsilon(1-\epsilon)u_1 + 2(1-\epsilon^2)u_2 + 2\epsilon(1-\epsilon)b, \end{gathered}$$

and hence

$$u_1 + u_7 + u_8 + 2(u_2+u_3+u_9) \equiv -3\epsilon^2 u_1 + 3u_2 - 3\epsilon^2 b,$$

it follows that the sum of the abelian parameter at a set of the $g_3{}^2$ is equal to $3\epsilon^2 u_1 - 3u_2 + 3\epsilon^2 b$. It is easily seen that the birational transformation

$$\bar{\gamma} : u' = \epsilon^2 u + \epsilon^2 u_1 - u_2 + \epsilon^2 b$$

of ψ into itself sends the $g_3{}^2$ of the line sections into the $g_3{}^2$ cut out by Φ'. Therefore $\bar{\gamma}T_4$ is involved by a collineation λ sending ψ into ψ_0. λ sends the points $1, 7, 8, 2, 3, 9$ into the points $1_0, 7_0, 8_0, 2_0, 3_0, 9_0$. For, the point 1 for example, goes by $\bar{\gamma}$ into the point $2\epsilon^2 u_1 - u_2 + \epsilon^2 b$ which by (7) is the sixth intersection of ψ with the conic $(1\,7\,2\,3\,9)$ which goes by T_4 into one of the points $1_0, 7_0, 8_0$, say 1_0, and in the same way 2 goes into 2_0 etc. The product λT_4^{-1} gives $\bar{\gamma}$ which is evidently a biquadratic transformation having the fundamental points of the direct and inverse transformation coincident, and hence there is a net of equianharmonic cubics on $(1\,7\,8\,2\,3\,9)$ from which follows that every pencil of type (m) is also of type (c).

Since in the mapping M the net of equianharmonic sections of F_3 go into equianharmonic quartics we conclude also that on 7 of the 9 base-points of a pencil of equianharmonic cubics containing only one triple there is a net of equianharmonic elliptic quartics.

THE SYMMETRIC (n, n)-CORRESPONDENCE AND SOME GEOMETRIC APPLICATIONS.

By Arnold Emch.

I. Introduction.

Two parameters λ and μ shall respectively determine two geometric entities of the same kind uniquely. For example, the values of either λ or μ may determine the tangents of a conic, or the planes of a pencil, or the osculating planes of a space cubic, or the quadrics of a pencil, etc. We then can set up an algebraic relation between λ and μ as the variables and study the properties of this relation. To these correspond in the geometric association of λ and μ definite geometric properties which by this method are often obtained in a surprisingly simple manner when compared with the arduous synthetic labors necessary to obtain the same result. This will be apparent in the applications to the Schur sextic with ∞^1 inscribed pentahedrons. This interesting space curve has received very little attention in the literature and deserves detailed study. The same is true of the Lüroth quartic with its ∞^1 inscribed pentalaterals and its connection with the Schur sextic and the attached octavic ruled surface.

II. The Symmetric (n, n)-Correspondence.

Although there is nothing essentially new in the algebraic exposition of this correspondence,* I shall give here a short account which will be needed subsequently.

Let $\lambda + \mu = x$, $\lambda\mu = y$, then a polynomial of degree n in x and y

$$(1) \qquad R(x, y) = 0.$$

represents such a correspondence. It depends evidently on $\frac{1}{2}n(n+3)$ constants and represents a general n-ic in the (x, y)-plane and is therefore of genus $\frac{1}{2}(n-1)(n-2)$. In the (λ, μ)-plane it is a curve of order $2n$ with the infinite points of the λ- and μ-axis as n-fold points, whose genus is $\frac{1}{2}(2n-1)(2n-2) - n(n-1) = (n-1)^2$. Geometrically, the correspondence between λ and μ may be obtained as follows: Associate λ and μ

* See A. B. Coble, "Multiple Binary Forms with Closure Property," *American Journal of Mathematics*, Vol. 43 (1921), pp. 1-19.

as parameters with the tangents of a fixed conic K. In the plane of K choose a general n-ic C_n as given by (1). From every point of C_n draw the two tangents to K and denote their parameters by λ and μ; then the algebraic relation between these parameters is precisely of the symmetric (n, n)-type. From this follows

THEOREM 1. *The pairs* $(\lambda, \mu) \equiv (\mu, \lambda)$ *of the symmetric* (n, n)*-correspondence are in birational correspondence with the points of a* C_n *of genus* $\frac{1}{2}(n-1)(n-2)$. There exists a $(1, 2)$-correspondence between C_n and C_{2n}. To every point of C_n correspond two points of C_{2n} which are symmetric with respect to the line $\lambda - \mu = 0$ in the (λ, μ)-plane.

In case of a "complete symmetry," as I define it, it must be true that for a generic value of $\lambda = \lambda_1$, we obtain n values for μ, $\mu = \lambda_2, \lambda_3, \cdots, \lambda_{n+1}$, such that every couple $(\lambda = \lambda_i, \mu = \lambda_k)$; $(i, k = 1, 2, \cdots, n+1;\ i \neq k)$, is also a pair of corresponding values (λ, μ), or a solution of (1). Geometrically these values determine two pencils of parallel lines

$$\lambda = \lambda_1, \lambda_2, \cdots, \lambda_{n+1}, \qquad \mu = \lambda_1, \lambda_2, \cdots, \lambda_{n+1}$$

which intersect in $(n+1)^2$ points; or, omitting the points on $\lambda - \mu = 0$, in $n(n+1)$ points of the curve C_{2n}. These form $\frac{1}{2}n(n+1)$ symmetric pairs (λ, μ); (μ, λ) with respect to the line $\lambda - \mu = 0$. Each of these pairs determines one of the $\frac{1}{2}n(n+3)$ constants of (1). Thus there are $\frac{1}{2}n(n+3) - \frac{1}{2}n(n+1) = n$ left to be determined, which can be done by the series of new couples $(\lambda, \mu) = (\lambda_{n+1}, \lambda_{n+2}), (\lambda_{n+1}, \lambda_{n+3}), \cdots, (\lambda_{n+1}, \lambda_{2n+1})$. Since the C_{2n} passes through the above $n(n+1)$ points it must necessarily have the form, after compounding it with $\lambda - \mu = 0$, and the line at infinity,

$$(2) \qquad C^*_{2n} = \prod_{i=1}^{n+1} (\lambda - \lambda_i) P(\lambda, \mu) + \prod_{i=1}^{n+1} (\mu - \lambda_i) Q(\lambda, \mu) = 0,$$

in which P and Q are of degree $n+1$. But since C_{2n} has the infinite points of the λ- and μ-axis as n-fold points, P and Q must be polynomials of degree $n+1$ in λ and μ alone, respectively, so that each can be resolved into $n+1$ linear factors

$$P = b \prod_{i=n+2}^{2n+2} (\mu - \lambda_i), \qquad Q = a \prod_{i=n+2}^{2n+2} (\lambda - \lambda_i).$$

As (2) must vanish identically, for $\lambda = \mu$, there must be $b = -a$, so that the relation is finally

$$(3) \quad \prod_{i=1}^{n+1} (\lambda - \lambda_i) \prod_{i=n+2}^{2n+2} (\mu - \lambda_i) - \prod_{i=1}^{n+1} (\mu - \lambda_i) \prod_{i=n+1}^{2n+2} (\lambda - \lambda_i) = 0.$$

After division by $\lambda-\mu$ and cancellation of $\lambda^{n+1}\mu^{n+1}$ this becomes the completely symmetric (n, n)-correspondence. Choosing

$$\prod_{i=1}^{n+1} (\mu-\lambda_i)/\prod_{i=n+2}^{2n+2} (\mu-\lambda_i)=k$$

as a parameter, (3) appears as an involution of order $n+1$ on the line λ, so that *"complete symmetry" is equivalent with the involutorial property.* (3) may be considered as the product of the two projective involutorial pencils, with t as a parameter:

$$\prod_{i=1}^{n+1} (\lambda-\lambda_i)+t\prod_{i=n+2}^{2n+2} (\lambda-\lambda_i)=0,$$

$$\prod_{i=1}^{n+1} (\mu-\lambda_i)+t\prod_{i=n+2}^{2n+2} (\mu-\lambda_i)=0.$$

For every value of t we obtained two pencils of $n+1$ parallel lines each, which intersect in $n(n+1)$ points of the curve C_{2n}. Thus, there are ∞^1 such sets. The result may be stated as

THEOREM 2. *The symmetrical involutorial (n, n)-correspondence between λ and μ is completely determined by two involutorial sets of $\frac{1}{2}n(n+1)$ couples (λ, μ) each. If the correspondence contains one involutorial set, then it contains ∞^1 such sets. One involutorial set and n other couples $(\lambda_{n+2}, \lambda_i)$ $(i=n+3, \cdots, 2n+2)$, determine the correspondence completely.*

III. THE PARAMETERS λ, μ ASSOCIATED WITH THE TANGENTS OF A CLASS-CONIC.

1. *The involutorial (n, n)-correspondence.* Again let

$$R(x, y)=0, \qquad x=\lambda+\mu, \qquad y=\lambda\mu, \tag{4}$$

represent such a correspondence, and λ the parameter which for every value determines a tangent

$$a\lambda^2+b\lambda+c=0 \tag{5}$$

of a class-conic K, in which a, b, c are linear forms of a plane $(x)=(x_1, x_2, x_3)$. Such a tangent is likewise

$$a\mu^2+b\mu+c=0. \tag{6}$$

For every couple (λ, μ) satisfying (4), (5) and (6) represent two tangents of K which intersect in a point P of a certain curve. To find the equation of this curve, λ and μ must be eliminated between (4), (5), (6). For this

purpose subtract (5) from (6) and divide the result by $\lambda - \mu$. This gives $ax + b = 0$. Again multiply (5) and (6) by μ^2 and λ^2 respectively, subtract and divide by $\lambda - \mu$, which gives $by + cx = 0$. Solving the two resulting equations, we get $x = -b/a$, $y = c/a$, which on substitution in (4) give an equation of degree n in x_1, x_2, x_3, representing a curve C_n of order n. As there are two tangents (λ), (μ) from every point of C_n to K, the n-ic is evidently in birational relation with the pairs $(\lambda, \mu) \equiv (\mu, \lambda)$ or with the n-ic $R(x, y)$ in the (x, y)-plane. The genus of C_n is therefore $\frac{1}{2}(n-1)(n-2)$. Moreover every tangent (λ) determines an involutorial $(n+1)$-lateral which is inscribed to the C_n, i. e., whose $\frac{1}{2}n(n+1)$ points of intersection lie on C_n. Hence

THEOREM 3. *The involutorial (n, n)-correspondence determines a curve of order n of genus $\frac{1}{2}(n-1)(n-2)$ with ∞^1 completely inscribed $(n+1)$-laterals, whose sides touch a fixed conic.* For $n = 2, 3, 4$, we have the following well-known results:

2. *Triangles inscribed and circumscribed to two conics respectively.* For $n = 2$, the involutorial correspondence determines a conic C_2 to which are inscribed ∞^1 triangles whose sides touch a given conic K.

3. *Complete quadrilaterals inscribed in a cubic.* The case $n = 3$ leads to an elliptic cubic C_3 with ∞^1 quadrilaterals circumscribed to a conic with their vertices on the fixed cubic. The result may be stated differently as

THEOREM 4. *Two sextuples of vertices formed by each of two distinct quadrilaterals circumscribed to a fixed conic determine a plane cubic uniquely. There are ∞^1 quadrilaterals inscribed to the cubic and circumscribed to the conic.*

4. *The Lüroth quartic.* When $n = 4$, we get a quartic of genus 3 with ∞^1 completely inscribed pentalaterals whose sides envelope a fixed conic. This quartic was found by Lüroth * in connection with the study of the Clebsch quartic,† which is characterized by the existence of an apolar conic, or its representability by the sum of the squares of five linear forms. These represent in every case the five sides of a pentalateral whose ten vertices generate the Lüroth quartic.

* "Einige Eigenschaften einer gewissen Gattung von Curven vierter Ordnung," *Mathematische Annalen*, Vol. 1 (1869), pp. 37-53.

† "Ueber Curven vierter Ordnung," *Crelle's Journal für reine und angewandte Mathematik*, Vol. 59 (1861), pp. 125-145.

Propositions of this and a similar sort have also been investigated in a different connection by W. F. Meyer.*

IV. The Involutorial (n, n)-Correspondence on a Developable Cubic.

1. *General case.* Assume a fixed developable cubic surface D with the likewise rational cubic edge of regression Γ and its tangents t. D is of order 4; i. e., through every generic line in space there pass four tangents t of Γ, and every t is the intersection of two infinitely close planes of D. This surface may be represented in the form

$$a\lambda^3 + b\lambda^2 + c\lambda + d = 0, \tag{7}$$

in which a, b, c, d are linear forms in $S_3(x_1, x_2, x_3, x_4)$. For every value of λ, (7) is one of the planes (7) of D. Likewise

$$a\mu^3 + b\mu^2 + c\mu + d = 0 \tag{8}$$

is such a plane. The parameters λ and μ are now subjected to the involutorial relation expressed by (1), respectively (3). Every couple (λ, μ) satisfying this relation determines two planes of D which intersect in a generatrix g of a certain ruled surface G, whose equation is again obtained by eliminating λ and μ between (1), (7) and (8). Setting, as before, $\lambda + \mu = x$, $\lambda\mu = y$, then by simple manipulations of (7) and (8) the equations result:

$$a(x^2 - y) + bx + c = 0, \tag{9}$$

$$by^2 + cxy + d(x^2 - y) = 0, \tag{10}$$

The elimination gives first an equation of degree $4n = n \cdot 2 + n \cdot 2$. Now (9) and (10) are satisfied by $x = - c/b$, $y = x^2 = (c/b)^2$, which solution does not satisfy (1). As a consequence the result of the elimination is reduced by $2n$, so that for the order of the ruled surface G proper, the number $2n$ is obtained. A generic plane (λ_1) determines an involutorial group of $n + 1$ planes (λ_1), $(\lambda_2), \cdots, (\lambda_{n+1})$ on D, so that on each of these, the remaining n cut out on n-lateral with $\frac{1}{2} \cdot n(n-1)$ points of intersection. In each of these planes there are no other points in which two planes of the involutorial set are concurrent. The points of the n-lateral in (λ_i), as λ_1 varies, describe therefore a space curve C_m of order $m = \frac{1}{2}n(n-1)$. Any two planes of the $(n+1)$-edron intersect in a generatrix g of G, and g is cut by $n - 1$ planes of the involutorial group and becomes thus an $(n-1)$-fold secant of the space curve. Moreover through

* *Apolarität und rationale Curven* (1883).

every point of C_m there are three $(n-1)$-fold secants, so that C_m is a triple curve of G. A plane (λ) of D cuts G in n lines g and a residual C_n to which the n-lateral (g) is inscribed. A generic plane p cuts G in a C_{2n} which at the $m=\frac{1}{2}n(n-1)$ intersections of C_m with p has triple points. The genus of C_{2n} is therefore

$$\pi=\tfrac{1}{2}(2n-1)(2n-2)-3\cdot\tfrac{1}{2}n(n-1)=\tfrac{1}{2}(n-1)(n-2).$$

Both C_{2n} and C_n are on the same ruled surface G and hence birationally related, so that C_n has also the genus π, and is therefore without singularities. The genus σ of C_m is easily found from the projection C_m of C_m from one of its generic points upon a generic plane. C_m has the order $\frac{1}{2}n(n-1)-1$ and has three $(n-2)$-fold points arising from the three $(n-1)$-fold secants of C_m through the center of projection. Thus

$$\sigma=\tfrac{1}{2}\left[\tfrac{1}{2}n(n-1)-2\right]\left[\tfrac{1}{2}n(n-1)-3\right]-3\cdot\tfrac{1}{2}(n-2)(n-3),$$

which reduces to

$$\sigma=1/8(n^4-2n^3-21n^2+70n-48).$$

A plane $(\lambda)=(\lambda_1)$ of D and an infinitely close plane (λ'_1) intersect in a tangent t of Γ, the curve of regression of D. The remaining planes of the involutorial group determined by (λ_1), cut (λ_1) in the n generatrices $g_2, g_3, \cdots, g_{n+1}$. These cut t in n points $P_2, P_3, \cdots, P_{n+1}$ which are on G as well as on D. Likewise to (λ'_1) correspond n planes which cut (λ'_1) in n generatrices $g'_2, g'_3, \cdots, g'_{n+1}$ which are infinitely close to $g_2, g_3, \cdots, g_{n+1}$, respectively, and which cut t in n points $P'_2, P'_3, \cdots, P'_{n+1}$, in the same order infinitely close to $P_2, P_3, \cdots, P_{n+1}$. Hence t touches G in n points $P_2, P_3, \cdots, P_{n+1}$. The plane (λ_1) cuts G in $g_2, g_3, \cdots, g_{n+1}$, and a residual C_n. This curve cuts t in n points which are precisely the points $P_2, P_3, \cdots, P_{n+1}$, so that these are the points of tangency of (λ_1) with G. From this follows that D as a surface of order 4 touches G along a curve of order $4n$, which counted twice is the complete intersection of D and G. In conclusion we have

THEOREM 5. *The $\frac{1}{2}n(n+1)$ lines of intersection g of the planes of a group of $n+1$ planes corresponding to the groups of an involutorial (n,n)-correspondence on a cubic surface D of class 3 generate a ruled surface G of order $2n$ with a space curve C_m of order $m=\frac{1}{2}n(n-1)$ and genus $\frac{1}{2}(n^4-2n^3-21n^2+70n-48)$ as a triple curve. Every generatrix of G is an $(n-1)$-fold secant of C_m. Every plane of D cuts G in n generatrices and a residual C_n. The residual n intersections of these generatrices with C_n*

are collinear on a line t which is a tangent of the cubic of regression of D. The latter surface is a complete contact surface of G and touches G along a curve of order $4n$.

2. *A sextic surface with a space cubic as a triple curve.* This surface G is obtained for $n = 3$ and was first investigated by E. Weyr * and also by the author † as an example of a more general class of surfaces.

The preceding theorem 5 for this case may be stated in a form found by A. Hurwitz ‡ in a remarkable study on problems of closure:

THEOREM 6. *The eight faces of any two tetrahedrons inscribed in a space cubic are osculating planes of another space cubic. Both curves are related suchwise that there are ∞^1 such tetrahedrons with the same property. Every point of the first curve is a vertex of just one of these tetrahedrons.*

The two curves are $C_m = C_3$ of genus $\sigma = 0$, and Γ. Each face of a tetrahedron inscribed to C_3 and circumscribed to Γ cuts G in three generatrices g and a residual elliptic cubic $C_n = C_3$. The g's cut C_3 in three residual points which are collinear on a tangent t of Γ. D touches G along a curve of order 12.

3. *The Schur sextic.* In 1881 F. Schur published a paper on curves and surfaces generated by collinear forms (Grundgebilde) in which, among other interesting results, he found a space sextic with remarkable properties. This is the curve $C_m = C_6$ of genus $\sigma = 3$, when $n = 4$ in the involution on the developable surface D of class 3. The five planes of an involutorial group form a pentahedron completely inscribed in the C_6; i. e., its 10 edges are trisecants and through each of the 10 vertices pass three trisecants. On account of the importance of this curve it is perhaps not superfluous to restate the general results for this particular case in the form of

THEOREM 7. *Among the space sextics C_6 of genus 3 there exists a class in which every curve admits of ∞^1 inscribed pentahedrons. The edges of these pentahedrons generate for every curve C_6 a ruled surface G of order 8 with C_6 as a triple curve. If a C_6 has one inscribed pentalateral, then it has an infinite number. The faces of the pentahedrons envelope a developable*

---

* "Ueber Flächen sechsten Grades mit einer dreifachen cubischen Curve," *Wiener Sitzungsberichte*, Vol. 85 (1882), pp. 513-525.

† "Ueber eine besondere Klasse von algebraischen Flächen," *Commentarii Mathematici Helvetici*, Vol. 2 (1930), pp. 99-115.

‡ "Ueber unendlich-vieldeutige geometrische Aufgaben, insbesondere über Schliessungsprobleme," *Mathematische Annalen*, Vol. 15 (1879), pp. 8-15.

surface D of class 3 and order 4 with a cubic curve of regression Γ. D is a contact surface of G. Every face of an inscribed pentahedron cuts G in a quadrilateral (g) and a residual quartic C_4. The sides of (g) cut C_4 in four residual points, collinear on a tangent t of Γ, so that (g) and t form an inscribed pentalateral of C_4. These residual curves are therefore Lüroth quartics. The manifold of Schur sextics is ∞^{20}.

The concluding statement may easily be verified. Two arbitrary planes (λ_1) and (λ'_1) of D cut every C_6 of the class. With (λ_1) we can associate four other planes of D to form a pentahedron Δ; likewise with (λ'_1) another distinct set of four planes of D forms a second pentahedron Δ'. The two Δ and Δ' determine a Schur sextic uniquely and may be chosen in $\infty^4 \cdot \infty^4 = \infty^8$ ways. On the other hand D is determined by six generic planes (λ). But on a given D six planes may be chosen in ∞^6 ways. The manifold of D's is therefore $\infty^{18} \cdot \infty^6 = \infty^{12}$, and consequently the manifold of Schur sextics $\infty^8 \cdot \infty^{12} = \infty^{20}$. It will be remembered that it is ∞^{24} for the general sextic of genus 3.

4. *The Schur sextic as a locus of vertices of cones.* In Schur's paper can also be found the proof that the sextic discussed above is also a "Kegelspitzenkurve," i. e., the locus of vertices of cones in a certain net of quadrics, or the Jacobian curve of this net. Now the manifold of such Jacobian curves is ∞^{21}. Putting the condition on such a curve that it shall lie on a six-point of a quadrilateral diminishes the manifold by 1, so that the manifold of these Jacobian curves is ∞^{20}, the same as that of the Schur sextics.

The fact that the Schur sextic is the Jacobian curve of a net of quadrics may easily be verified by the following proof which may be given in outline. Take any of the inscribed pentahedrons of a C_6. It contains 5 tetrahedrons. On each face of such a tetrahedron there lies a quadrilateral of trisecants with a diagonal triangle. Thus on each of the tetrahedrons there are 12 points which lie on a quadric Q, and every inscribed pentahedron determines 5 quadrics which may be shown to be linearly independent and form a linear four dimensional system $\sum_{i=1}^{5} a_i Q_i = 0$, with respect to which the pairs of vertices of every quadrilateral of trisecants are apolar. Two such systems of quadrics based on two distinct inscribed pentalaterals have a net of quadrics in common, whose fundamental curve in the involutorial cubic transformation determined by the net is identical with the Schur sextic. Thus

THEOREM 8. *The Jacobian curve of a net of quadrics is a Schur sextic when it is on the six points of a plane quadrilateral.*

A HYPERSURFACE OF ORDER 2^{r-1} IN r-SPACE.

By B. C. Wong.

1. The equation

$$(1) \qquad x_1^{1/2} \pm x_2^{1/2} \pm \cdots \pm x^{1/2}_{r+1} = 0,$$

when rationalized, is of degree 2^{r-1} and therefore represents a hypersurface $V^{2^{r-1}}_{r-1}$ of order 2^{r-1} in r-space. If $r = 2$ and 3, we have a conic and a Steiner's quartic surface respectively. In this paper we propose to study the hypersurface for $r > 3$. We shall first obtain the different multiple varieties on the hypersurface, then give a few general remarks concerning its 3-space and plane sections. We shall, finally, consider its representation upon a hyperplane.

2. For the purpose of obtaining the different multiple varieties on $V^{2^{r-1}}_{r-1}$, we find it convenient to group all the terms in the left-hand member of (1) into ν groups. Letting the k-th group contain t_k terms, we write the equation in the form

$$(2) \quad \sum_1^{t_1} x_{i_1}^{1/2} \pm \sum_1^{t_2} x_{t_1+i_2} \pm \sum_1^{t_3} x^{1/2}_{t_1+t_2+i_3} \pm \cdots \pm \sum_1^{t_\nu} x^{1/2}_{t_1+t_2+\cdots+t_{\nu-1}+i_\nu} = 0$$

where

$$(3) \qquad t_1 + t_2 + \cdots + t_\nu = r + 1,$$

or

$$\sum_1^{\nu} \Phi_k = 0$$

where

$$\Phi_k = \sum_1^{t_k} x^{1/2}_{t_1+t_2+\cdots+t_{k-1}+i_k}.$$

3. Consider one of the $(t_k - 1)$-spaces, S_{t_k-1}, of the coördinate simplex, Δ_r, and let its equations be given by all the $r + 1 - t_k$ x's not contained in Φ_k equal to zero. This S_{t_k-1} is tangent to $V^{2^{r-1}}_{r-1}$ along a $V^{2^{t_k-1}}_{t_k-1}$ 2^{r-t_k}-uply. The equations of this $V^{2^{t_k-1}}_{t_k-1}$ of contact are the $r + 1 - t_k$ equations just mentioned and $\Phi_k = 0$. The nature of $V^{2^{t_k-1}}_{t_k-1}$ is identical with that of $V^{2^{r-1}}_{r-1}$ for $r = t_k - 1$. By allowing t_k to take on successively the values $2, 3, 4, \cdots, r$, we find that the hypersurface is touched, respectively, by each of the $\binom{r+1}{2}$ edges of Δ_r in 2^{r-2} coincident points, by each of the $\binom{r+1}{3}$ planes of Δ_r along 2^{r-3} coincident conics, by each of the $\binom{r+1}{4}$ 3-spaces of Δ_r along 2^{r-4} coincident Steiner's quartic surfaces, $\cdots$, by each of the $r + 1$ hyperplane faces of Δ_r along a $V^{2^{r-2}}_{r-2}$.

4. Now the ν equations

$$(4) \qquad \Phi_1 = 0, \quad \Phi_2 = 0, \cdots, \quad \Phi_\nu = 0$$

represent ν hypersurfaces

$$V_{r-1}^{2^{t_1-1}}, \quad V_{r-1}^{2^{t_2-1}}, \cdots, \quad V_{r-1}^{2^{t_\nu-1}}$$

respectively. The hypersurface $V_{r-1}^{2^{t_k-1}}$ given by $\Phi_k = 0$ is the locus of the ∞^{t_k-2} S_{r+1-t_k}'s all passing through the S_{r-t_k} of Δ_r whose equations are obtained by equating to zero all the t_k x's contained in Φ_k and meeting the opposite S_{t_k-1} of Δ_r in the points of the variety $V_{t_k-2}^{2^{t_k-2}}$ of contact between S_{t_k-1} and $V_{r-1}^{2^{r-1}}$. The equations (4) taken simultaneously represent a $V_{r-\nu}^{2^{t_1+t_2+\cdots+t_\nu-2\nu}} \equiv V_{r-\nu}^{2^{r+1-2\nu}}$ which is the variety of intersection of the ν hypersurfaces and is the locus of the $\infty^{r+1-2\nu}$ $S_{\nu-1}$'s incident with all the varieties of contact

$$V_{t_1-1}^{2^{t_1-2}}, \quad V_{t_2-2}^{2^{t_2-2}}, \cdots, \quad V_{t_\nu-2}^{2^{t_\nu-2}}.$$

This variety $V_{r-\nu}^{2^{r+1-2\nu}}$ is of multiplicity $2^{\nu-1}$ on $V_{r-1}^{2^{r-1}}$. As the $r+1$ terms in the left-hand member of (1) can be grouped into ν groups as indicated in (2) in

$$N = (1/\nu!)\sum (r+1)!/(t_1!\,t_2!\cdots t_\nu!)$$

ways for all integral values of the t's greater than unity and satisfying relation (3), there are N $2^{\nu-1}$-fold varieties of dimension $r-\nu$ and order $2^{r+1-2\nu}$ on $V_{r-1}^{2^{r-1}}$, constituting a composite $2^{\nu-1}$-fold variety of dimension $r-\nu$ and order

$$M = (2^{r+1-2\nu}/\nu!)\sum (r+1)!/(t_1!\,t_2!\cdots t_\nu!).$$

5. For a given of value of ν, the N $2^{\nu-1}$-fold $V_{r-\nu}^{2^{r+1-2\nu}}$'s are of different types, depending on the values of the t's. Of course, for $\nu = 1$, we have $V_{r-1}^{2^{r-1}}$ itself. For $\nu = 2$, we find that the entire composite double variety on $V_{r-1}^{2^{r-1}}$ is of order $M = 2^{r-3}(2^r - r - 2)$ and is composed of $N = 2^r - r - 2$ ruled varieties all of order 2^{r-3}. These component varieties are distributed as follows: For t_1 or $t_2 = 2$, we have $\binom{r+1}{2} = \binom{r+1}{r-1}$ $(r-2)$-dimensional cones of order $r-3$ each having for vertex the point where an edge of Δ_r touches $V_{r-1}^{2^{r-1}}$ and standing on the $V_{r-3}^{2^{r-3}}$ of contact in the opposite S_{r-2} of Δ_r. Similarly, for t_1 or $t_2 = 3$, we have $\binom{r+1}{3} = \binom{r+1}{r-2}$ $V_{r-3}^{2^{r-3}}$'s each of which has for generators the lines incident with the conic of contact in a plane of Δ_r and the $V_{r-4}^{2^{r-4}}$ of contact in the opposite S_{r-3} of Δ_r. Continuing in this manner by assigning all possible values to t_1, t_2 satisfying the relation $t_1 + t_2 = r+1$, we shall soon have obtained all the double varieties on the hypersurface $V_{r-1}^{2^{r-1}}$.

6. We may proceed in a similar manner for the other values of ν, but we shall be content with an illustration. For $r=5$, the double variety V_3^{100} of order 100 on V_4^{16} is composed of $N=25$ V_3^{4}'s of which 15 are 3-dimensional cones each having its vertex on an edge of Δ_5 and standing on the Steiner's quartic surface of contact in the opposite S_3 of Δ_5 and 10 are ruled quartic varieties each being the locus of the ∞^2 lines incident with the two conics of contact lying in opposite planes of Δ_5. When $\nu=3$, $t_1=t_2=t_3=2$, V_4^{16} has 15 quadruple planes each being determined by the points of contact on non-adjacent edges of the coördinate simplex.

7. On the composite double variety on $V_{r-1}^{2^{r-1}}$ is a pinch variety which is of dimension $r-3$ and order $2^{r-4}r(r+1)$ and is composite, being composed of the $\binom{r+1}{2}$ $V_{r-3}^{2^{r-3}}$'s along which the S_{r-2}'s of Δ_r are tangent to $V_{r-1}^{2^{r-1}}$ doubly. Thus, the Steiner's quartic surface has six pinch points and the V_3^8 in S_4 has 10 pinch conics.

8. On the hypersurface $V_{r-1}^{2^{r-1}}$ there are numerous other multiple varieties whose multiplicities are not powers of 2. We shall here indicate a simple method by means of which these varieties may be obtained in any given case. Returning to the equation, (1), of the hypersurface, we find that the signs in its left-hand member may be combined in 2^r ways. Then, the left-hand member of the equation, after it is rationalized, is the product of the 2^r factors each with a different combination of signs. Now any point whose coördinates make q of these 2^r factors vanish is a q-fold point on the hypersurface. This method also enables us to find the $2^{\nu-1}$-fold varieties already mentioned above.

9. As an illustration, consider the V_3^8 in S_4. It has 10 double quadric cones K^2_{ij} whose equations are

$$x_i : x_j : x_k : x_l : x_m = u^2 : u^2 : v^2 : (v-w)^2 : w^2$$

$[i, j, k, l, m=1, 2, \cdots, 5]$. The right-hand members of these equations make the two factors

$$x_i^{1/2}-x_j^{1/2}+x_k^{1/2}-x_l^{1/2}-x_m^{1/2}, \qquad x_i^{1/2}-x_j^{1/2}-x_k^{1/2}+x_l^{1/2}+x_m^{1/2}$$

vanish. Now let $v=u$ and we have

$$x_i : x_j : x_k : x_l : x_m = u^2 : u^2 : u^2 : (u-w)^2 : w^2$$

which are the equations of 10 conics C^2_{ijk}, each being common to the three K^2_{ij}, K^2_{jk}, K^2_{ki} and lying in the plane $x_i=x_j=x_k=0$. As the coördinates

of the points on these conics make the three factors

$$-x_i^{1/2}+x_j^{1/2}+x_k^{1/2}-x_l^{1/2}-x_m^{1/2}, \qquad x_i^{1/2}-x_j^{1/2}+x_k^{1/2}-x_l^{1/2}-x_m^{1/2},$$
$$x_i^{1/2}+x_j^{1/2}-x_k^{1/2}-x_l^{1/2}-x_m^{1/2}$$

equal to zero, the conics are triple curves on $V_3{}^8$. Now putting $w=+2u$, we obtain five points whose coördinate $(1:1:1:1:4)$ make the four factors

$$+x_i^{1/2}-x_j^{1/2}-x_k^{1/2}-x_l^{1/2}+x_m^{1/2}, \qquad -x_i^{1/2}+x_j^{1/2}-x_k^{1/2}-x_l^{1/2}+x_m^{1/2},$$
$$-x_i^{1/2}-x_j^{1/2}+x_k^{1/2}-x_l^{1/2}+x_m^{1/2}, \qquad -x_i^{1/2}-x_j^{1/2}-x_k^{1/2}+x_l^{1/2}+x_m^{1/2}$$

vanish. These points are, therefore, quadruple points on the hypersurface $V_3{}^8$.

10. From what has been said concerning the singular varieties on the hypersurface $V_{r-1}^{2^{r-1}}$ it is not difficult to describe the singularities on any of its sections. In particular, the surface, $F^{2^{r-1}}$, in which a general S_3 meets it contains a composite double curve consisting of 2^r-r-2 components all of order 2^{r-3}. Of these component curves $r(r+1)/2$ are plane curves. Denote these by $C_{ij}^{2^{r-3}}$. They are the intersections of the given S_3 and the $r(r+1)/2$ $(r-2)$-dimensional double cones of order 2^{r-3} on $V_{r-1}^{2^{r-1}}$. They are all of deficiency $2^{r-5}(r-6)+1$, having $2^{r-5}(2^{r-2}-r)$ nodes. Each of these nodes is the common node of three curves $C_{ij}^{2^{r-3}}$, $C_{kl}^{2^{r-3}}$, $C_{mn}^{2^{r-3}}$ and is a quadruple point on $F^{2^{r-1}}$. There are in all

$$M_{\nu=3}=(2^{r-5}/3!)\sum(r+1)!/t_1!\,t_2!\,t_3!$$

such quadruple points where $t_1+t_2+t_3=r+1$. Now any three curves $C_{ij}^{2^{r-3}}$, $C_{ik}^{2^{r-3}}$, $C_{jk}^{2^{r-3}}$ which do not have a node in common meet in 2^{r-3} points on the line common to their planes. These points are triple points on $F^{2^{r-1}}$ and there are in all $r(r^2-1)2^{r-3}/6$ $[r>3]$ such triple points.

11. A general plane section, $c^{2^{r-1}}$, of $V_{r-1}^{2^{r-1}}$ has $2^{r-3}(2^r-r-2)$ nodes and $r+1$ r-fold tangents. Each of the latter is the intersection of the plane of $c^{2^{r-1}}$ and an S_{r-1} of Δ_r. As each r-fold tangent counts as $r(r-1)/2$ double tangents, $c^{2^{r-1}}$ has $2^{r-2}(2^{r-3}r^2-5r+8)-r(r^2-1)/2$ double tangents besides. The curve $c^{2^{r-1}}$ is of class $2^{r-2}r$ and is of deficiency $2^{r-3}(r-4)+1$.

12. All the previous results can be readily obtained from the following representation of the hypersurface upon a hyperplane R_{r-1}. Let $y_1, y_2, \cdots, y_{r+1}$ be the homogeneous point coördinates of R_{r-1} where

$$y_1+y_2+\cdots+y_{r+1}\equiv 0$$

and let the equation

(5) $$u_1y_1^2 + u_2y_2^2 + \cdots + u_{r+1}y^2_{r+1} = 0$$

define an ∞^r-system $|Q|$ of $(r-2)$-dimensional quadric varieties in R_{r-1}. Taking the members of $|Q|$ as fundamental varieties, we can represent the hypersurface $V^{2^{r-1}}_{r-1}$ upon R_{r-1}, the formulas of representation being

$$x_1 : x_2 : \cdots : x_{r+1} = y_1^2 : y_2^2 : \cdots y^2_{r+1}.$$

If we eliminate the y's from these equations, the result is equation (1). Each of the $(r-2)$-spaces y_i taken doubly is a member of $|Q|$.

13. By means of this representation we see that the hypersurface we are studying is of order 2^{r-1} as any $r-1$ general quadric varieties of $|Q|$ intersect in 2^{r-1} points. We now show that it is of class r. Consider the Jacobian varieties $J^{(1)}_{r-2}, J^{(2)}_{r-2}, \cdots, J^{(r-1)}_{r-2}$, all of order r, of any $r-1$ linear ∞^{r-1}-systems of $|Q|$. They all pass $h-1$ times through each of the $\binom{r+1}{h}$ S_{r-h-1}'s in which the $r+1$ S_{r-2}'s : y_i intersect h by h. Thus, each of the $\binom{r+1}{2}$ S_{r-3}'s in which the y's intersect two by two, each of the $\binom{r+1}{3}$ S_{r-4}'s in which the y's intersect three by three, $\cdots$, each of the $\binom{r+1}{r-2}$ lines in which the y's intersect $r-2$ by $r-2$, and finally, each of the $\binom{r+1}{r-1}$ points in which the y's intersect $r-1$ by $r-1$ are simple, double, $\cdots$, $(r-3)$-fold, and $(r-2)$-fold, respectively, on all the Jacobian varieties. Any two of these varieties, say $J^{(1)}_{r-2}, J^{(2)}_{r-2}$, intersect in a $V^{\binom{r}{2}}_{r-3}$ besides the $\binom{r+1}{2}$ S_{r-3}'s mentioned above. Now this $V^{\binom{r}{2}}_{r-3}$ intersects any third one, say $J^{(3)}_{r-2}$, in a $V^{\binom{r}{3}}_{r-4}$ besides the $\binom{r+1}{3}$ double S_{r-4}'s also mentioned above. Continuing in this manner, we find that the number of free points common to all the $r-1$ Jacobian varieties is $\binom{r}{r-1} = r$. Hence, the class of $V^{2^{r-1}}_{r-1}$ is r.

14. Of course, this fact can be easily derived from the equation of the hypersurface. Or, it may be seen from the fact that a general pencil of quadric varieties of $|Q|$ contains r conic varieties. Any general hyperplane S_{r-1} with the equation

$$u_1x_1 + u_2x_2 + \cdots + u_{r+1}x_{r+1} = 0$$

meets $V^{2^{r-1}}_{r-1}$ in a $V^{2^{r-1}}_{r-2}$ which corresponds to a quadric variety Q of $|Q|$. S_{r-1} will be a tangent hyperplane if Q is singular. The discriminant of equation (5) equated to zero, is, after dividing by $u_1u_2 \cdots u_{r+1}$,

$$1/u_1 + 1/u_2 + \cdots + 1/u_{r+1} = 0.$$

This is the equation of $V^{2^{r-1}}_{r-1}$ in hyperplane coördinates. Interpreting u_i as

point coördinates of S_r, we can regard this equation as representing a hypersurface $V^r{}_{r-1}$ of order r and class 2^{r-1}, reciprocal to $V^{2^{r-1}}_{r-1}$. $V^r{}_{r-1}$ has on it $r+1$ $(r-1)$-fold points, $\binom{r+1}{2}$ $(r-2)$-fold lines joining the $(r-1)$-fold points two by two, $\binom{r+1}{3}$ $(r-3)$-fold planes joining the $(r-1)$-fold points three by three, $\cdots$, $\binom{r+1}{r-1}$ simple S_{r-2}'s joining the $(r-1)$-fold points $r-1$ by $r-1$.

15. It is evident that $V^r{}_{r-1}$ is representable upon R_{r-1} by means of the r-ic varieties of the type $J^{(1)}_{r-2}$, $J^{(2)}_{r-2}$, $\cdots$. Attention is here called to the involutorial r-ic transformation in S_r effected by means of r hyperquadric surfaces. To a point P we make correspond the point P' of intersection of its polar hyperplanes with respect to them. If P (or P') describes an S_{r-1} or in particular R_{r-1}, then P' (or P) describes an r-ic hypersurface. This hypersurface will be the $V^r{}_{r-1}$ reciprocal to $V^{2^{r-1}}_{r-1}$ if the r hyperquadric surfaces have a self-polar simplex, Δ_r, in common. The intersection of Δ_r and R_{r-1} is the configuration, Δ'_r, formed by the $r+1$ $(r-2)$-spaces y_i. The quartic varieties in R_{r-1} coresponding to the ∞^r hyperplane sections of $V^r{}_{r-1}$ pass through the $\binom{r+1}{r-1}$ $(r-3)$-spaces of Δ'_r and form an ∞^r-system. The Jacobian variety of a general linear ∞^{r-1}-family of this system is of order $r(r-1)$ but it is composed of the $r+1$ $(r-2)$-spaces y_i each taken $r-2$ times and a quadric variety whose equation is of the form (5). $r-1$ such Jacobian varieties yield $r-1$ quadric varieties having 2^{r-1} points in common, and, thus, we are led back to the representation of the hypersurface $V^{2^{r-1}}_{r-1}$ upon R_{r-1}.

16. If we take all the members of the $\infty^{(r+1)(r-1)/2}$-system of $(r-2)$-dimensional quadric varieties in R_{r-1} for fundamental varieties of representation, we can set up a one-to-one correspondence between the points of an $(r-1)$-dimensional variety of order 2^{r-1} in an $S_{(r+2)(r-1)/2}$ and the points of R_{r-1}. The projection upon an S_r of this variety from a general $S_{(r^2-r-4)/2}$ in $S_{(r+2)(r-1)/2}$ as center of projection is the $V^{2^{r-1}}_{r-1}$ we are investigating.

17. A general plane section $c^{2^{r-1}}$ of $V^{2^{r-1}}_{r-1}$ has for image the curve $C^{2^{r-1}}$ in which $r-2$ general quadric varieties of the system $|Q|$ in R_{r-1} intersect. If $C^{2^{r-2}}$ is projected on to an S_3, the projection will have $2^{r-3}(2^{r-2}-r+1)$ apparent double points. Hence $C^{2^{r-2}}$ is of deficiency $2^{r-3}(r-4)+1$ and, therefore, $c^{2^{r-1}}$ is of deficiency $2^{r-3}(r-4)+1$ as has already been mentioned.

THE UNIVERSITY OF CALIFORNIA.

TWO INVOLUTORIAL TRANSFORMATIONS, OF ORDERS 11 AND 9, ASSOCIATED WITH NULL RECIPROCITIES.

By VIRGIL SNYDER AND HAZEL E. SCHOONMAKER.

1. *Statement of the problem.* An interesting transformation may be defined as follows: Take four mutually skew lines a_i in space and a fixed point A not lying on any of them. Any plane through A cuts these lines in four points A_i. Then there exists in that plane a point O such that the lines OA, OA_i are projective with five given lines of a pencil. If the plane is allowed to describe the bundle A, the point O will describe a cubic surface * passing through A and containing the four lines a_i. This cubic surface is uniquely fixed. The analytic procedure will show that the converse is true. Hence we now have a $(1,1)$ correspondence between the points of a cubic surface and the planes of a bundle through a point on it such that each point lies in its corresponding plane.

A $(1,1)$ correspondence between two spaces such that the points of one space correspond to the planes of the other, each point lying in its corresponding plane, is called a null reciprocity. Montesano † has shown that such a transformation can be considered as the product of a birational transformation and a correlation, and that every birational transformation can be obtained as the product of a null reciprocity and a correlation.

The transformation described above is not good for all space. It will become so if, instead of keeping A fixed, we let A describe a line u. We then obtain a pencil of cubic surfaces $\psi + a\theta = 0$.

Any point P in space uniquely fixes a cubic of the pencil and a point A on u. The line AP meets the cubic associated with A in the point P', image of P. This transformation is involutorial.

2. *The parameters of the five lines OA_i.* Following analytically the steps outlined above, we first get the parameters of the five lines OA_i. Let the four skew lines be

$$a_1 \equiv x_3 = 0,\ x_1 + x_2 + x_4 = 0; \qquad a_2 \equiv x_4 = 0,\ x_1 - x_2 + x_3 = 0;$$
$$a_3 \equiv x_3 - x_4 = 0,\ x_1 + 2x_2 + x_3 = 0; \quad a_4 \equiv x_3 - 2x_4 = 0,\ 2x_1 + x_2 + x_4 = 0.$$

* Proved by Sturm, "Ueber correlative oder reciproke Bündel," *Mathematische Annalen*, Vol. 12 (1877), § 38.

† "Sulle reciprocità birazionali nulle dello spazio," *Rendiconti della reale Accademia dei Lincei*, Vol. 4 (1888), pp. 583-590.

Let $A \equiv (0, 0, a, 1)$ be any point on $u \equiv x_1 = 0$, $x_2 = 0$. Then any plane through A is

$$\pi \equiv c_1x_1 + c_2x_2 + c_3x_3 - ac_3x_4 = 0.$$

The plane $\pi = 0$ meets a_i in the points A, A_i. Let $O \equiv (y_1, y_2, y_3, y_4)$ and consider the coördinates of any point on OA_i. Cut these lines by the plane $x_1 = x_3$. This will give five points on the line $\pi = 0$, $x_1 = x_3$. Denote these points by B_i. For lack of space we omit the coördinates of these points.

The coördinates of any point on the line $\pi = 0$, $x_1 = x_3$ can be expressed linearly in terms of those of B, B_1. Then for a properly chosen h and k, $hB + kB_1 = B_i$. If we take B for $(1, 0)$, B_1 for $(0, 1)$, and put $h/k = \lambda_i$ for the parameter of B_i and hence of OA_i, then by solving the proper equations we find these parameters to be

$$\lambda_2 = \frac{c_1(y_1 + y_2 - y_3 + y_4) + c_2(y_1 + y_2 + y_3 + y_4) + 2c_3y_3}{ay_4 - y_1 + y_2 - y_3}$$

$$\lambda_3 = \frac{c_1(y_3 - 2y_1 - 2y_2 - 2y_4) + c_2(y_1 + y_2 + y_4) + c_3y_3(a-1)}{ay_1 + 2ay_2 + ay_4 - y_1 - 2y_2 - y_3}$$

$$\lambda_4 = \frac{c_1(2y_1 + 2y_2 + 2y_4) - c_2(4y_1 + 4y_2 + 4y_4 - y_3) + c_3y_3(a-1)}{4y_1 - 2ay_1 - ay_2 + y_2 + y_3 - ay_4}$$

3. *Equations of the pencil of cubic surfaces.* Let λ be the cross-ration of the first three and fourth lines; and λ' the cross-ratio of the first three and fifth lines. Then $\lambda = \lambda_3/\lambda_2$, $\lambda' = \lambda_4/\lambda_2$. Substituting the values of λ_i and remembering that O lies in $\pi = 0$, we see that O satisfies the following equations

$$\begin{aligned} c_1\alpha_1 + c_2\alpha_2 + c_3\alpha_3 &= 0, \\ c_1\beta_1 + c_2\beta_2 + c_3\beta_3 &= 0, \\ c_1y_1 + c_2y_2 + c_3(y_3 - ay_4) &= 0, \end{aligned}$$

where α_i are functions of (y) and λ obtained by simplifying $\lambda = \lambda_3/\lambda_2$, and β_i are functions of (y) and λ'. Again we omit numerical details. For these equations to have a solution the determinant must vanish. Hence $\Delta = 0$ is the locus of 0. This determinant reduces to

$$\begin{vmatrix} \lambda(-2y_1-4y_2-2y_3)+2y_1+4y_2-y_3+3y_4 & y_3(-y_1+y_2-y_3)+y_4(3y_1+3y_2+3y_4) \\ \lambda'(8y_1+4y_2+2y_3)-8y_1-4y_2+y_3-6y_4 & y_3(-y_1+y_2-y_3)+y_4(-6y_1-6y_2-6y_4+3y_3) \end{vmatrix}$$

$$+a\begin{vmatrix} \lambda(2y_1+4y_2+2y_4)+y_4-y_3 & y_3(-y_1+y_2-y_3)+y_4(3y_1+3y_2+3y_4) \\ \lambda'(-4y_1-2y_2-2y_4)+2y_4-y_3 & y_3(-y_1+y_2-y_3)+y_4(-6y_1-6y_2+3y_3-6y_4) \end{vmatrix}$$

$$\equiv \psi + a\theta = 0.$$

4. *The complete intersection of two surfaces of the pencil.* The pencil $\psi + a\theta = 0$ contains A but not u, the locus of A, and it contains the four skew lines. Since the surfaces are cubics and contain four skew lines, they must contain their transversals, which are found to be

$$t_1 \equiv x_3 = 0,\ x_4 = 0;\quad t_2 \equiv 3x_1 + x_3 = 0,\ 2x_1 + x_2 + x_4 = 0.$$

These may be obtained by building the quadric surface containing a_1, a_2, a_3 and finding the points in which it is cut by a_4. The lines of the other regulus passing through these two points are the desired transversals. These lines do not meet u.

The residual intersection of any two cubics of the pencil is a space cubic curve c_3, $p = 0$ and of rank 4, meeting each a_i in two points, t_j in no points, and u in two points. If we find the points in which u meets $\psi = 0$ and $\theta = 0$ we find that u is a bisecant of c_3. The other properties of c_3 can be obtained by examining the map of a cubic surface.

5. *The equations of the transformation.* Any point on AP has coördinates of the form

$$\rho y'_1 = \sigma y_1,\ \rho y'_2 = \sigma y_2,\ \rho y'_3 = \sigma y_3 - \tau\psi,\ \rho y'_4 = \sigma y_4 + \theta\tau.$$

The line AP meets $F_3 = 0$ in the point P', image of P. We wish to determine σ and τ so that (y') lies on F_3. Substituting the above values in the expression for F_3, we obtain

$$F_3 \equiv \theta(y)\psi(y') - \psi(y)\theta(y') = 0.$$

Expanding by Taylor's series and recalling that A and (y) are on the surface, we find that

$$\sigma = \psi(y)\theta_2(y, \psi, \theta) - \theta(y)\psi_2(y, \psi, \theta),$$
$$\tau = \psi(y)\theta_1(y, \psi, \theta) - \theta(y)\psi_1(y, \psi, \theta),$$

where θ_i is the i-th polar of A as to θ, etc. More explicit values of σ and τ are rather complicated and we omit them here.

6. Order of the transformation and multiplicity of the fundamental curves. It is seen that σ is of order 10 and τ of order 8, hence the transformation is of order 11. Also σ is cubic in ψ and θ is quadratic in these quantities. It follows that the intersections of $\psi = 0$, $\theta = 0$, namely, $4a_i$, $2t_j$, c_3 are triple on the surfaces of the web. If A is any point on $u = 0$, it is found that the coördinates of A satisfy $\sigma = 0$, $\tau = 0$ and the first partial derivatives vanish for this point. Moreover,

$$\frac{\partial^2\sigma}{\partial y_i\partial y_k} = \psi(A)\frac{\partial^2\tau}{\partial y_i\partial y_k} \neq 0, \quad \text{and} \quad \frac{\partial^2 y'_i}{\partial y_i\partial y_k} = \frac{\partial^2\sigma}{\partial y_i\partial y_k} - \psi(A)\frac{\partial^2\tau}{\partial y_i\partial y_k} = 0.$$

Hence the surfaces $\sigma = 0$ and $\tau = 0$ have contact along a double line and this line is triple on the surfaces of the web.

In addition to these curves there are twenty parasitic lines. Eight of them, denoted by p_m, meet u, a_i, a_j, a_k; four, denoted by q, meet u, c_8 and the two transversals; and eight, denoted by g_{ij}, meet u, c_8, a_i, and t. To obtain p_4, for example, consider the F_2 containing a_1, a_2, and a_3, and the points in which it is met by u. Through each of these points is a line meeting a_1, a_2, a_3. Now consider the F_2 containing t_1, t_2 and u. This is met by c_8 in four points not on u. Through each of these is a line q. If we cut the pencil of cubics by a plane through a_i and t_j the residual is a line. For one cubic surface this line passes through the point in which u meets the plane as well as through the point in which c_8 meets the plane.

All these lines meet u. Let K be the point in which one of them g meets u. The cubic surface of the pencil belonging to K contains g, and by the transformation any point of g is transformed into the whole line g and into nothing else. Hence each of these lines is a simple fundamental line of the second kind.

7. *Images of the fundamental curves.* It is obvious that the image of $u = 0$ is $\sigma = 0$, and that any plane through u is invariant. Hence in any such plane π we have an involution. The plane π contains u and seven points $4A_i$, $2T_j$, $C_{8,1}$, the other two points $C_{8,2}$, and $C_{8,8}$ in which c_8 meets the plane being on u. The surface $\tau = 0$ is the surface of invariant points and its section by π is C_6: $4A_i^2 2T_j C^2_{8,1}$. Hence the involution in the plane is of the Geiser type. In this involution $A_i \sim C_8$: $A_i^2 3A_j 2T_k C_{3,1}$. If we revolve the plane about u, the Geiser involution will generate our space involution and hence the image of a_i is A_i: $a_i^2 3a_k 2t_j c_8 u$. Similarly $t_j \sim T_j$: $4a_i t_j^2 t_k c_8 u$. Since c_8 has two points on u its image is Γ_6: $4a_i 2t_j c_8^2 u^3$. Each of these surfaces contain some of the parasitic lines.

8. *Conclusion.* We find then that the transformation is given by

$$s_1 \sim s_{11}:\ 4a_i^3 2t_i^3 c_8^3 u_8 8p_i 4q 8g_{ik}.$$

The jacobian is composed of

$$\begin{aligned} A_i &: \ a_i^2 3a_k 2t_j c_8 u 8p_i 4g_{ij}, \\ T_j &: \ 4a_i t_j^2 t_k c_8 u 4q 4g_{ij}, \\ \Gamma_6 &: \ 4a_i 2t_j c_8^2 u^3 4q 8g_{ij}, \\ \sigma_{10} &: \ 4a_i^3 2t_j^3 c_8^3 u_2 8p_i 4q 8g_{ij}, \end{aligned}$$

conjugates of a_i, t_j, c_3 and u. The surface of invariant points is

$$\tau:\quad 4a_i{}^2 2t_j{}^2 c_3{}^2 u_2 8p_i 4q 8g_{ij}.$$

This I_{11} contains two independent linear parameters λ, λ', hence there are really defined ∞^2 transformations. Any two distinct elements (λ, λ'), (λ'', λ''') generate an infinite discontinuous group of Cremona transformations, each if which contains the fundamental elements u, $4a_i$, $2t_j$; the residual space cubics form a congruence having u and $4a_i$ for bisecants which has been studied by Godeaux * and by J. de Vries.†

9. *The second transformation.* Another interesting transformation may be defined as follows: Take the lines u and $4a_i$ as before. A plane α through u meets each a_i in A_i. A point P in α defines a cross-ratio λ determined by the four ordered lines PA_i. The locus of P as λ remains constant and α turns about u is a quartic surface F_4: $u^2 4a_i$. In its equation, λ appears linearly, hence as λ varies, a pencil of quartic surfaces $F_4 \equiv \psi + \lambda\theta = 0$ is defined. By the method above we obtain

$$\begin{aligned}
\psi \equiv\; & x_2{}^2(x_2 - x_3 + 3x_4)(x_1 + 2x_2 - x_3 + 2x_4) \\
& + x_1x_2\,[(x_2 - x_3 + 3x_4)(x_1 + 2x_2 + x_4) \\
& + (-2x_1 + x_2 - 2x_3 + 3x_4)(x_1 + 2x_2 - x_3 + 2x_4)] \\
& + x_1{}^2(-2x_1 + x_2 - 2x_3 + 3x_4)(x_1 + 3x_2 + x_4), \\
\theta \equiv\; & x_2{}^2(4x_1 + 2x_2 + 2x_4)(2x_1 - 2x_2 + 2x_3 - 3x_4) \\
& + x_1x_2\,[(4x_1 + 2x_2 + 2x_4)(+x_1 - x_2 + x_3) \\
& + (4x_1 + 2x_2 - x_3 + 4x_4)(2x_1 - 2x_2 + 2x_3 - 3x_4) \\
& + x_1{}^2(4x_1 + 2x_2 - x_3 + 4x_4)(x_1 + 2x_2 + x_4).
\end{aligned}$$

Now take a second line u' not meeting any of the given ones nor lying on any surface of the pencil. A point (y) determines a surface of the pencil $F_4(y)$. The transversal of u, u' through (y) meets $F_4(y)$ in one residual point (y'). The transformation $y \sim y'$ is birational and involutorial.

Let u' be the line $x_1 - x_4 = 0$, $x_2 - 2x_3 = 0$. The plane determined by u and (y) meets u' in $U' \equiv (2y_1, 2y_2, y_2, 2y_1)$. Any point on the line $U'(y)$ has coördinates of the form

$$\rho x_1 = (\sigma + 2\tau)y_1,\ \rho x_2 = (\sigma + 2\tau)y_2,\ \rho x_3 = \sigma y_3 + \tau y_2,\ \rho x_4 = \sigma y_4 + 2\tau y_1,$$

and

$$F_4(y) \equiv \psi(x)\theta(y) - \psi(y)\theta(x) = 0.$$

---

* "Sur une congruence linéo-linéaire de cubiques gauches, *Bulletins de la Classe des Sciences, Académie Royale de Belgique, 1809*, p. 531.

† The congruence of twisted cubics that cut five lines twice," *Proceedings Koninklijke Akademie van Wetenscappen te Amsterdam*, Vol. 31 (1928), pp. 454-458.

For the point (y') on this line we have

$$\tau = -2\,[\theta(y)\psi(\bar{y}, y) - \psi(y)\theta(\bar{y}, y)],$$
$$\sigma = \theta(y)\psi(\bar{y}) - \theta(\bar{y})\psi(y),$$

wherein $(\bar{y}) \equiv U'$, and $\psi(\bar{y}, y)$ is the polar of U' as to the quadric coefficients of x_1^2, $2x_1x_2$, x_2^2 in ψ, etc.

10. *The fundamental elements.* The transformation is of order 9. In $\sigma = 0$, u appears to order 6, and in $\tau = 0$, to order 5. Since τ is multiplied by y_1 or y_2 in each equation of the transformation the conjugates of the planes of space contain u^6.

From the definition of the transformation, every plane through u remains invariant. In each plane of this pencil the transformation is perspective Jonquières, of order 3, having U' for double point and $4A_i$ for simple fundamental points. Every conic of the pencil $A_1 \cdots A_4$ is transformed into itself except that one passing through U' which has the point U' for its conjugate. Hence u' is double on I_9. The locus of this conic as α turns about u' is $\sigma = 0$.

The line a_i is simple on I_9, since the conjugate of any point A_i on it is the line $U'A_i$. The locus of this line is the quadric $uu'a_i \equiv H_i$.

The lines meeting $ua_1a_2a_3 = p_4$, etc. are basis lines of the pencil of quartics, since u is a double line. The plane up_i meets u' in T'. Three points A_k lie on p_i and A_i does not. Consider any point (y) in this plane. The line joining it to A_i meets p_i in a point which determines λ, and the line $T'(y)$ meets p_i in its conjugate (y'). Thus the conjugate of p_i is the plane up_i. For one value of λ, the line $(y)A_i$ passes through T'. These lines are parasitic, or simple fundamental lines of the second kind.

Finally, consider the transversals of u, u', a_i, $a_k \equiv q_{ik}$. The plane u, q_{ik} meets u' in Q' on q_{ik}, hence each line q_{ik} is a simple fundamental line of the second kind. This completes the configuration of fundamental elements. The conjugate of u is $\sigma + 2\tau = 0$, and $\tau = 0$ is the locus of invariant points. The surfaces $\sigma = 0$, and $\tau = 0$ touch each other along the line u. Every plane through u' is invariant; in each plane is a perspective Jonquières I_7: $U'^6 4A_i 8Q_k$.

The table of characteristics has the form

$$s_1 \sim s_9 : u^6u'^2 4a_i 8p_i 12q_{ik} 8t_i,$$
$$u \sim U_8 : u^5u'^2 4a_i 8p_i 12q_{ik} 8t_i,$$
$$u' \sim U'_8 : u^6u' 4a_i 8p_i 12q_{ik} 8t_i,$$
$$a_i \sim H_{2i} : uu'a_i 6q_{ik} 2t_i,$$
$$p_i \sim \pi_i : up_it_i,$$
$$J_{32} \equiv U_8U'_8 4H_2,\ 8\pi_i.$$

SOME PROPERTIES OF THE FUNDAMENTAL CURVES OF A BIRATIONAL TRANSFORMATION IN SPACE.

By Marian M. Torrey.

Given a web of homaloidal surfaces $|\phi|$ in space (x), homogeneous of degree n in x_1, x_2, x_3, x_4, and forming a regular system, that is, one having no fixed tangent planes at any isolated fundamental point, or at any point of a fundamental curve. A $(1,1)$ correspondence is defined between spaces (x) and (y) by the equations

$$y_1 : y_2 : y_3 : y_4 = \phi_1 : \phi_2 : \phi_3 : \phi_4,$$
$$x_1 : x_2 : x_3 : x_4 = \psi_1 : \psi_2 : \psi_3 : \psi_4,$$

where the ψ's are homogeneous of degree m in y_1, y_2, y_3, y_4, and also form a regular system. To a line of (x) corresponds a curve C'_n in (y), and to a line of (y) corresponds a curve C_m in (x). Suppose the web $|\phi|$ has a fundamental curve u of the first kind, order ν, multiplicity k. Then it is known * that if all surfaces of the web tangent to a fixed plane π at a point P of u are considered, they form a net which corresponds to a bundle of planes in (y), with vertex at a point P'. As π turns about the line through P tangent to u, the point P' generates a curve v_P of order k, which is the image of P, and as P moves along u, v_P generates a surface of degree equal to the number of variable points in which any C_m meets u. Every plane through P corresponds to a ψ passing simply through the curve v_P.† The planar neighborhood π of P corresponds to the spatial neighborhood of P'.

The object of the present discussion is to study this correspondence of directions at P and P' in more detail. It will be shown that a direction l at P in π corresponds to a planar neighborhood at P' in π'_l, a plane through the line tangent to v_P at P': as l rotates about P in π, π'_l revolves about the tangent to v_P. A direction at P' in π'_l determines a point in the second neighborhood of P on a curve C_m tangent to l at P: stated a little differently,

* D. Montesano, "Su la teoria generale delle corrispondenze birazionali fra i punti dello spazio," *Atti della R. Accademia di Napoli*, Ser. 2, Vol. 17 (1927), No. 8, p. 4.

† L. Godeaux, "Sur les courbes fondamentales des transformations birationelles de l'espace," *Bulletins de l' Académie royale de Belgique*, Ser. 5, Vol. 15 (1929), pp. 317-318.

any C_m tangent to l corresponds to a line of π'_l; the osculating plane of C_m at P corresponds to a surface tangent to π'_l at P', and having the line corresponding to C_m for an inflexional tangent.

It is clear first that the surface ψ corresponding to π, which may be called ψ_π, has a double point at P', for since every curve C_m which is the intersection of two surfaces of the net has one more intersection with π than with any other plane through P, all lines through P' have one more intersection with ψ_π than with other ψ's containing v_P. Hence any pencil of planes with axis l through P in π corresponds to a pencil of surfaces through v_P, all having a fixed tangent plane at P', namely π'_l, which passes through the line tangent to v_P. As l rotates about P, remaining in π, the plane π'_l changes position, always containing the tangent to v_P at P'.

There are ∞^1 pencils of surfaces ϕ tangent to π at P, each having its curve of intersection C_m tangent to l at P. The condition for this is that every plane through l should cut two surfaces of the pencil defining C_m in two plane curves, which in addition to having k-fold points at P, and the contact along one branch which determines the net, have contact of the second order along this branch, that is, in terms of cartesian coördinates, have equal second derivatives at P. This is a linear condition on the parameters of the net, reducing the net to a pencil. Furthermore the condition varies with the particular value imposed on the second derivative, giving ∞^1 such pencils.

Every plane through l cuts such a C_m in two points at P, and one plane through l, the osculating plane of C_m, meets it in three points at P. In space (y), therefore, every surface tangent to π'_l touches the line corresponding to such a C_m, so that this line must be in the plane π'_l. One surface tangent to π'_l at P', the one corresponding to the osculating plane of C_m, has this line for its inflexional tangent.

For an accurate picture of this relationship between the lines of π'_l and the surfaces ψ tangent to π'_l, note first that each line of π through P corresponds to a composite curve in (y) made up of v_P and a variable curve C'_{n-k}, these two curves intersecting at P'. A generic plane through P' meets this composite curve in two points at P'; any plane of the pencil with axis tangent to v_P meets the curve in three points at P'; the plane tangent to both v_P and C'_{n-k} meets the curve in four points at P'. Hence any line of π through P meets a general surface of the net tangent to π in $k+1$ points at P, and a surface corresponding to a plane through the tangent to v_P in $k+2$ points. That is, such a surface has two sheets tangent to π, or if v_P is a straight line so that u is simple, a surface corresponding to a plane through v_P has

a double point at P, since in this case such a surface must belong to every net defined by a fixed tangent plane at P. Furthermore the surface which corresponds to the plane π'_l tangent to both v_P and C'_{n-k}, is met by the corresponding l in $k+3$ points at P, that is, such an l must be an inflexional tangent for one of the sheets of this surface tangent to π, or in the case when u is simple, l must be tangent to the surface.

Now consider the curves C_m tangent to a particular l of π and corresponding to the lines of π'_l through P'. All such curves lie on the surface corresponding to π'_l, which may be called ϕ_l. Any curve osculating a particular plane α of the pencil through l must have three consecutive points in the plane α at P, and these points must also be on ϕ_l. Since ϕ_l has two sheets tangent to π, there are two curves on which the points can approach P, remaining in α and on ϕ_l, and one of these curves osculates l, since one sheet of ϕ_l osculates l. Hence two C_m osculate each α, one osculating l, and the other osculating the second branch tangent to l of the section of ϕ_l by α.

If another C_m osculated this plane, it would have two points consecutive to P in common with one of the above curves (C_m). But two C_m can intersect in only one variable point.

The C_m osculating l osculates every plane through l. Hence each plane α through l is osculated by two and only two curves C_m through P, one peculiar to α, and one osculating every plane through l. It follows that one line of π'_l is an inflexional tangent for every surface of the pencil tangent to π'_l; this line is also tangent to the proper image of l, a C'_{n-k}. The second inflexional tangent in π'_l varies with the surface of the pencil.

When k equals one, a similar argument shows that there is only one curve C_m osculating each plane through l, since only one branch of the section of ϕ_l by a plane through l is tangent to l. This curve varies with the plane through l. The pencil of surfaces in (y) tangent to π'_l has one variable inflexional tangent at P', the other being the line v_P.

Hence there is a $(1, 1)$ correspondence between directions at P in π, and planar neighborhoods at P' defined by the planes passing through the tangent to v_P, and there is a $(1, 1)$ correspondence between neighborhoods of the second order at P defined by the osculating planes of curves C_m tangent to the same line of π, and directions at P' in the corresponding planar neighborhood.

Now suppose the $|\phi|$ have a fundamental curve of the second kind, u, order ν, multiplicity $k\nu'$, which corresponds to a fundamental curve of the second kind v on the web $|\psi|$, order ν', multiplicity $k\nu$, where k is a positive integer. Then if the condition is imposed that some surfaces of the web

in (x) be tangent at P, a point of u, to an arbitrary plane π_1 through the tangent line to u, a net of surfaces results all tangent at P to π_1 and to $k-1$ other planes $\pi_2, \cdots, \pi_k$, and these k planes form an involution, any one determining the other $k-1$.* Furthermore the net so determined has k fixed tangent planes at every point of u. This net corresponds to a bundle of planes with vertex at a point of v, say P'. As the set of k planes through P varies, the point P' generates v. A similar situation exists for nets $|\psi|$ tangent to a fixed plane at a point of v. All ψ tangent to one fixed plane at a point of v are tangent to $k-1$ others, and as the set of k planes varies, the point remaining fixed, the vertex of the corresponding bundle of planes in (x) generates the curve u.

At a generic point P of u, determine a net $|\phi|$ tangent at P to π_1. This fixes a set of planes $\pi_1, \pi_2, \cdots, \pi_k$, whose equations are

$$\sum_{j=1}^{4} a_{ij}x_j = 0; \qquad (i = 1, 2, \cdots, k).$$

Call the vertex of the corresponding bundle of planes in (y) P', and consider the particular set of k planes through the tangent to v at P' which determines the point P of u. It will be shown that the planes $\pi_1, \cdots, \pi_k$, can be paired uniquely with the k planes at P' in such a way that directions in π_i at P correspond to directions in π'_i at P', and vice versa. In other words, whatever the value of k, any tangent plane to u uniquely determines a tangent plane to v, such that the two planar neighborhoods correspond in the transformation. It follows that as P varies on u, the net $|\phi|$ being fixed, P' remains fixed, and the k planes π'_i, corresponding to the planes tangent to $|\phi|$ at P, generate the pencil whose axis is the tangent to v at P'. The spatial neighborhood of P' is thus shown to correspond to ∞^1 planar neighborhoods defined by the tangent planes to the fixed net $|\phi|$ at points of u.

The plane π_1 is tangent to all surfaces of the net $|\phi|$ corresponding to the bundle of planes, vertex P'. Hence any line l in π_1 through P meets any surface of the net in $k\nu' + 1$ points at P. The curve C'_n corresponding to such a line is composite. One component is a curve $C'_{n-k\nu'}$, the proper image of l, and the other is the curve v counted k times.* Since l meets any surface of the net in $n - k\nu' - 1$ variable points, $C'_{n-k\nu'}$ meets every plane of the bundle at P'. It must now be shown that all $C'_{n-k\nu'}$ corre-

* D. Montesano, "Sulla teoria generale delle corrispondenze birazionali dello spazio," *R. Accademia dei Lincei Atti. Rendiconti*, Ser. 5, Vol. 30 (1921), pp. 447-451.

sponding to lines of π_1, and therefore lying on the surface $\sum_{j=1}^{4} a_{1j}\psi_j = 0$, are tangent to one and only one of the k planes through P'.

The line l meets a generic plane of (x) in one variable point, and a generic plane of the bundle, vertex P, at P. The curve $C'_{n-kv'}$ meets any ψ of the web in one variable point, and in kv' fixed points at P'. It must therefore meet any surface of the net corresponding to the bundle of planes, vertex P, in $kv'+1$ points at P'. This can happen only if it is tangent at P' to one of the planes tangent to the net $|\psi|$. It cannot be tangent to more than one, for if tangent to i of the k planes, $C'_{n-kv'}$ would have $kv'+i$ points of intersection at P with every surface of the net, which would mean i points of intersection at P for l and every plane of the bundle. Furthermore, as l varies in π_1, turning about P, $C'_{n-kv'}$ varies continuously, always passing through P', so that it is not possible to have some of these $C'_{n-kv'}$ tangent to one of the k planes π'_i at P', and others tangent to another. Call the plane to which these $C'_{n-kv'}$ are tangent π'_1, and let its equation be

$$\sum_{j=1}^{4} b_{1j}y_j = 0.$$

Since it has been shown that $C'_{n-kv'}$ has two intersections at P' with π'_1, l must have $kv'+2$ intersections at P with the surface corresponding to π'_1, namely $\sum_{j=1}^{4} b_{1j}\phi_j = 0$, and since l is any line of π_1, the section of this surface by π_1 has a $kv'+2$-fold point at P.

It is shown in a similar way that the section of the surface $\sum_{j=1}^{4} a_{1j}\psi_j = 0$ by π'_1 has a $kv+2$-fold point at P', while the section of any other surface of the net by π'_1 has a $kv+1$-fold point at P'. For a line l' through P' in π'_1 corresponds to a curve C_{m-kv} which passes through P and is tangent to one of the k common tangent planes of the net $|\phi|$. That this plane is π_1 is evident from the fact that l' meets a C'_n on $\sum_{j=1}^{4} a_{1j}\psi_j = 0$, which is composed of v taken k times and the proper image $C_{n-kv'}$, in two consecutive points at P'. Hence the curve C_m, the complete image of l', meets l, the image of C'_n, in two consecutive points at P, so that l is a line in the plane tangent to u and C_{m-kv}, the components of C_m. But l is any line through P in $\pi_1 \equiv \sum_{j=1}^{4} a_{1j}x_j = 0$, and hence C_{m-kv} is tangent to π_1. The remainder of the argument is identical.

Furthermore the correspondence must be unique. For suppose lines of

a second plane π_2 in (x) corresponded to $C'_{n-k\nu'}$ which also were tangent to the plane π'_1. Then lines in the plane π'_1 would correspond to curves $C_{m-k\nu}$ tangent to both π_1 and π_2, and this has been shown to be impossible. Hence π_2 is paired with another of the k planes in (y) which we call π'_2. Proceeding thus, the k planes in (y) can be named $\pi'_1, \pi'_2, \cdots, \pi'_k$.

We now consider whether there is a correspondence between directions through P in π_i, and directions through P' in π'_i. Take a pencil of planes in the bundle, vertex P, with axis in π_i,

$$\sum_{j=1}^{4} c_j x_j + \lambda \sum_{j=1}^{4} a_{ij} x_j = 0,$$

where $\sum c_j x_j = 0$ is satisfied by the coördinates of P. In (y) space the corresponding pencil of surfaces has the equation

$$\sum_{j=1}^{4} c_j \psi_j + \lambda \sum_{j=1}^{4} a_{ij} \psi_j = 0,$$

all members of the pencil being tangent to the planes $\pi'_1, \cdots, \pi'_k$. Take a section of this pencil by the plane π'_i. A pencil of plane curves results with a $k\nu + 1$-fold point at P'. All tangents to this pencil at P' are fixed, since the section of $\sum_{j=1}^{4} a_{ij}\psi_j = 0$ by π'_i has been shown to have a $k\nu + 2$-fold point at P', and therefore this equation has vanishing derivatives of order $k\nu + 1$. This was to have been expected, since the curve of intersection of surfaces of the pencil has, in addition to the fundamental curve v, $k\nu$-fold for the system, two other branches through P' tangent to π'_i, namely a curve v_i consecutive to v, and the curve $C'_{n-k\nu'}$, the proper image of a line of π_i through P. This means that surfaces of a pencil in (y) corresponding to planes of a pencil with axis in π_i, have contact of the first order at P' along the sheets tangent to $\pi'_1, \pi'_2, \cdots, \pi'_{i-1}, \pi_{i+1}, \cdots, \pi'_k$, and contact of the second order at P' along the sheets tangent to π'_i.* The necessary though not sufficient condition for such contact is that the $k\nu + 1$ inflexional tangents to the surfaces at P' in π'_i shall be fixed for the pencil. $k\nu - 1$ tangents coincide with the tangent line to v, leaving two fixed directions in π'_i distinct from this line. A section of this pencil of surfaces by any other plane π'_j will not give fixed tangents at P' in that plane, and any pencil of planes with axis not in π_i, corresponds to a pencil of surfaces which does not have fixed tangents at P' in π'_i.

* E. Pascal, *Repertorium der Höheren Mathematik*, Vol. II₂, p. 654.

Similarly if the bundle of planes in (y) is considered, and a pencil of the bundle determined with axis in π'_i, a pencil of curves is found in π_i with two fixed tangents at P distinct from the tangent to u. However it cannot be said that the direction in π'_i given by the axis of the pencil of planes corresponds to the directions given by the two fixed inflexional tangents in π_i of the corresponding pencil of surfaces. The correspondence in directions at P and P' in the planes π_i and π'_i must be considered in two ways. One approaches P in π_i either along a curve C_{m-kv}, image of a line l' of π'_i, or along the tangent l to that curve. These two directions, apparently one, give different approaches in π'_i to P'. If P is approached along C_{m-kv}, P' is approached along a line l' of π'_i which is one of the inflexional tangents of the pencil of surfaces whose curve of intersection, $C'_{n-kv'}$, corresponds to l. If P is approached along l, P' is approached along $C'_{n-kv'}$, which is not in general tangent to l'.

If neither u nor v is a straight line, there are two curves C_{m-kv} tangent to each l of π_i, and two curves $C'_{n-kv'}$ tangent to each l' of π'_i. However, if one of the fundamental curves, say v, is a straight line, there is only one curve C_{m-kv} tangent to each l, but two curves $C'_{n-kv'}$ tangent to each l'; if both u and v are straight lines, there is one C_{m-kv} tangent to each l, and one $C'_{n-kv'}$ tangent to each l'.

A simple example of a fundamental curve of the second kind for which $k=2$ is given by Miss Hudson.* u and v are both straight lines, and the transformation, a cubic involution, is defined by the equations

$$x_1 : x_2 : x_3 : x_4 = y_1y_4^2 : y_2y_3^2 : y_3y_4^2 : y_3^2y_4.$$

v is the line $y_3 = y_4 = 0$. The system is not perfectly regular, since it has fixed tangent planes at the points $(1,0,0,0)$ and $(0,1,0,0)$, but these points may be avoided. As all known transformations which have fundamental curves of the second kind with $k>1$ are products of transformations for which $k=1$, so in this case the cubic is the product of the two quadratics

$$x_1 : x_2 : x_3 : x_4 = z_1z_3 : z_2z_4 : z_3^2 : z_3z_4,$$
$$z_1 : z_2 : z_3 : z_4 = y_1y_4 : y_2y_3 : y_3y_4 : y_3^2.$$

Consider a point on v, $(\xi,\eta,0,0)$. If the condition is imposed that at this point the plane $my_3 = y_4$ shall be tangent to the web

$$ay_1y_4^2 + by_2y_3^2 + cy_3y_4^2 + dy_3^2y_4 = 0,$$

* H. Hudson, *Cremona Transformations in Plane and Space*, Cambridge University Press (1927).

a second tangent plane is fixed, namely $my_3 = -y_4$, and the condition $a = -\eta b/\xi m^2$ is imposed on a and b. Hence the web becomes the net

$$b(\eta y_1 y_4^2 - \xi m^2 y_2 y_3^2) + c y_3 y_4^2 + d y_3^2 y_4 = 0,$$

which corresponds to the bundle of planes with vertex at $(\xi m^2, \eta, 0, 0)$.

Similarly if we take the web of surfaces in (x), the tangent planes at $(\xi m^2, \eta, 0, 0)$ are given by the equation

$$b\eta x_3^2 + a\xi m^2 x_4^2 = 0.$$

If one tangent plane is taken as $x_3 = mx_4$, the other is determined as $x_3 = -mx_4$, and $a = -\eta b/\xi$, so that the web is again reduced to a net

$$b(\eta x_1 x_4^2 - \xi x_2 x_3^2) + c x_3 x_4^2 + d x_3^2 x_4 = 0,$$

which corresponds to the bundle of planes with vertex at $(\xi, \eta, 0, 0)$. Any line through $(\xi, \eta, 0, 0)$ in the plane $my_3 = y_4$ of (y) has for image, in addition to the line $x_3 = x_4 = 0$ counted twice, a line through $(\xi m^2, \eta, 0, 0)$ in the plane $x_3 = mx_4$. This property of the transformation was noted by Miss Hudson in another connection.* In this simple case the curve $C_{m-k\nu}$ referred to above is a straight line; hence the problem of distinguishing between the direction at P in π_i as defined by $C_{m-k\nu}$ and by the tangent to $C_{m-k\nu}$ does not arise.

A simple non-involutorial transformation where u and v are not both straight lines can be obtained by applying a quadratic transformation to the above cubic. Thus if we combine

$$x_1 : x_2 : x_3 : x_4 = z_1 z_4^2 : z_2 z_3^2 : z_3 z_4^2 : z_3^2 z_4$$

with $$z_1 : z_2 : z_3 : z_4 = y_1 g : y_2 g : y_3 g : f,$$

where g is linear and f quadratic in $y_1, \cdots, y_4$, and in addition $f = 0$ is satisfied by the coördinates $(0, 0, 0, 1)$, there results a quintic

$$x_1 : x_2 : x_3 : x_4 = y_1 f^2 : y_2 y_3^2 g^2 : y_3 f^2 : y_3^2 g f$$

with the double conic $y_3 = f = 0$ fundamental of the second kind. The plane $g = 0$ cannot in this case be chosen as $y_4 = 0$, since that gives contact conditions all along the fundamental curve in (x) space. In any case, the web has fixed tangent planes at the points $y_2 = y_3 = f = 0$ and $y_1 = y_3 = f = 0$ which are distinct from $(0, 0, 0, 1)$, and at the points $y_3 = g = f = 0$, and these points must be avoided in applying the theory.

* H. Hudson, *loc. cit.*, p. 271.

It is easily seen that the inverse of the quintic is a sextic transformation which has the four-fold line $x_3 = x_4 = 0$, image of the fundamental curve $y_3 = f = 0$.

If a tangent plane π_i is fixed at a point P of the conic $y_3 = f = 0$, there is determined a tangent plane π'_i at a point P' of the line $x_3 = x_4 = 0$. A line l in π_i through P corresponds to a curve C'_3 tangent to π'_i at P'. A curve of the system tangent to l at P corresponds to a line in π'_i, which is an inflexional tangent common to the pencil of surfaces defining the C'_3 above, but this inflexional tangent of the surfaces is not tangent to C'_3: in other words, the line l and the curve tangent to l define a common direction at P, but their images define two distinct directions at P'.

The essential difference between the two cases arises from the fact that tangency to a fixed plane at a point P of a fundamental curve of the second kind entails tangency at every point of the curve, while for fundamental curves of the first kind this is not true. Hence there is contact of the second order for certain pencils of surfaces if P is on a fundamental curve of the second kind, and contact of the first order if the curve is of the first kind; hence also one, or at most two, curves of the system tangent to each line through P if u is of the second kind, and ∞^1 curves of the system tangent to each line through P if u is of the first kind.

ON SURFACES POSSESSING A NET OF PLANE ISOTHERMALLY-CONJUGATE CURVES.*

By C. A. Nelson.

1. *Introduction.* Let $y^{(k)} = y^{(k)}(u,v)$, $z^{(k)} = z^{(k)}(u,v)$, $(k = 1, 2, 3, 4)$ be the homogeneous coördinates of the two focal surfaces S_y, S_z—which we assume to be distinct—of any congruence, Γ, of lines. Let u and v be the parameters of those curves on S_y and S_z along which the developables of the congruence touch the focal surfaces. Then the congruence may be characterized, projectively, by a completely integrable system of partial differential equations of the form

$$\text{(D)} \qquad \begin{aligned} y_v &= mz, \qquad z_u = ny, \\ y_{uu} &= ay + bz + cy_u + dz_v, \\ z_{vv} &= a'y + b'z + c'y_u + d'z_v, \end{aligned}$$

in which the coefficients are analytic functions of u and v and the subscripts denote differention.†

The integrability conditions of system (D) are

$$\text{(I)} \qquad \begin{aligned} &c = f_u,\ d' = f_v,\ b = -d_v - df_v,\ a' = -c'_u - c'f_u,\ mn - c'd = f_{uv}, \\ &m_{uu} + d_{vv} + df_{vv} + d_vf_v - f_um_u = ma + db', \\ &n_{vv} + c'_{uu} + c'f_{uu} + c'_uf_u - f_vn_v = c'a + nb', \\ &2m_un + mn_u = a_v + f_umn + a'd, \\ &m_vn + 2mn_v = b'_u + f_vmn + bc', \end{aligned}$$

where f may be any analytic function of u and v.

The form of the equations (D) is unchanged by the group of transformations

$$\text{(T)} \qquad y = \lambda(u)\bar{y}, \quad z = \mu(v)\bar{z}, \quad \bar{u} = \phi(u), \quad \bar{v} = \psi(v),$$

λ, μ, ϕ and ψ being arbitrary functions. Although the form of (D) is unaltered, the coefficients themselves are changed. Among the combinations of these coefficients which are relatively invariant we cite

* Presented to the Society, December 27, 1928.

† E. J. Wilczynski, "Sur la théorie générale des congruences," *Mémoires publiés par la Classe des Sciences de l'Académie Royale de Belgique*, Deuxième serie, tome 3 (1911). This memoir will be referred to as the Brussels Paper.

$$(1.1) \qquad c', \; d, \; m, \; d_1 = d\left(c'd - \frac{\partial^2}{\partial u \partial v}\log d\right), \quad m_1 = m\left(mn - \frac{\partial^2}{\partial u \partial v}\log m\right).$$

Three of the covariants are

$$(1.2) \qquad \rho = y_u - (m_u/m)y, \quad \sigma = z_v - (n_v/n)z, \quad \epsilon = \rho_u - (\partial/\partial u)\log m_1 \cdot \rho.$$

The covariant σ, when z is replaced by $z^{(k)}$, defines the minus first Laplace transform of S_z. The surfaces S_σ and S_z are the focal surfaces of the minus first transform of Γ. Similarly, S_ρ and S_ϵ are the focal surfaces of the plus first transform of Γ.

If we wish to concentrate upon the focal surface S_y rather than upon the congruence Γ, we obtain a characterizing system of equations by eliminating z from (D). The result is

$$(\mathrm{Y}) \qquad \begin{aligned} y_{uu} &= (d/m)y_{vv} + f_u y_u - (d/m)(f_v + m_v/m + d_v/d)y_v + ay, \\ y_{uv} &= (m_u/m)y_{vv} + mny. \end{aligned}$$

The parametric curves upon S_y form a conjugate system. When it happens that this system has equal invariants and consists entirely of plane curves, the surface S_y is greatly restricted. It is the purpose of this paper to investigate the implications of these restrictions.

2. *Analytic formulation.* It is well known that if the $u =$ const. curves are plane the envelope of the tangents to these curves is developable or degenerate. The invariantive condition for this is *

$$(2.1) \qquad c' = 0.$$

Similarly, the $v =$ const. curves are plane if, and only if,

$$(2.2) \qquad d_1 = 0.$$

Finally, the conjugate system has equal invariants when, and only when,

$$(2.3) \qquad (\partial^2/\partial u \partial v)\log m = 0. \dagger$$

It should be remarked that the above conditions are also necessary and sufficient for a plane isothermally-conjugate net on S_y.‡

* Brussels Paper, p. 28. G. M. Green, "Projective Differential Geometry of One-parameter Families of Space Curves, and Conjugate Nets on a Curved Surface," *American Journal of Mathematics*, Vol. 37 (1915), p. 236.

† E. J. Wilczynski, "The General Theory of Congruences," *Transactions of the American Mathematical Society*, Vol. 16 (1915), p. 319. This paper will be cited as Congruence Paper.

‡ Congruence Paper, p. 322.

A comparison of the conditions (2. 1), (2. 2), (2. 3) with (1. 1) indicates that they may be replaced by

$$(2.4)\qquad c' = 0, \quad (\partial^2/\partial u \partial v) \log d = 0, \quad (\partial^2/\partial u \partial v) \log m = 0.$$

We shall assume throughout that $d \neq 0$, $m \neq 0$. Geometrically this means that the surface S_y is neither developable nor degenerate.*

Since no transformation (T) disturbs the form of (D), we seek a transformation that will simplify the conditions (2. 4). Now †

$$(2.5)\qquad \bar{m} = (\mu/\lambda)(1/\psi_v) m, \quad \bar{d} = (\mu/\lambda)(\psi/\phi_u^2) d,$$

so that if we write, from (2. 4), $d = U(u) \cdot V(v)$ and $m = U_1(u) \cdot V_1(v)$ the transformation for which

$$\lambda = U_1, \quad \mu = 1/(VV_1)^{1/2}, \quad \phi_u = (U/U_1)^{1/2}, \quad \psi_v = (V_1/V)^{1/2}$$

reduces both $\bar{m}$ and $\bar{d}$ to unity. Having made this transformation the conditions (2. 4) may be replaced by

$$(2.6)\qquad c' = 0, \quad d = m = 1.$$

The most general transformation not violating (2. 6) is given by

$$(2.7)\qquad \lambda = k_1, \quad \mu = k_1 k_2, \quad \phi_u = \psi_v = k_3,$$

where k_1, k_2 are constants.

The above values for c', d, m greatly simplify the integrability conditions. They become

$$(2.8)\qquad c = f_u, \quad d' = f_v, \quad b = -f_v, \quad a' = 0, \quad n = f_{uv}, \quad b' = f_{vv} - a,$$

$$(2.9)\qquad n_{vv} - f_v n_v = nb', \quad n_u = a_v + f_u n, \quad 2n_v = b'_u + f_v n.$$

The expressions (2. 6) and (2. 8) give all of the coefficients of (D) in terms of a, f, and their derivatives. The expressions (2. 9) represent conditions upon these two coefficients. Eliminating b' and n from (2. 9) we obtain

$$(2.10)\qquad a_v = f_{uuv} - f_u f_{uv}, \quad a_u = f_v f_{uv} - f_{uvv}, \quad a_{uv} = f_{uv} a.$$

We may sum up the results in the theorem: *Let a non-developable, non-degenerate surface S_y be referred to a conjugate system of plane curves with equal invariants. Then the congruence of tangents to the $u =$ constant curves may be characterized, projectively, by*

---

* Brussels Paper, p. 28.

† Brussels Paper, pp. 20, 23.

$$(\Delta)\qquad \begin{aligned} y_v &= z, \quad z_u = f_{uv} y,\\ y_{uu} &= ay - f_v z + f_u y_u + z_v,\\ z_{vv} &= (f_{vv} - a)z + f_v z_v, \end{aligned}$$

where a, f satisfy the conditions (2.10).

3. *An alternate form for* (2.10). We introduce the quantities

$$(3.1)\qquad h = f_{uv}, \quad h_1 = f_{uv} - (\partial^2/\partial u \partial v) \log f_{uv}.$$

Solving the first two equations of (2.10) for f_{uuv}, f_{uvv} and differentiating, we find

$$(3.2)\qquad f_{uuvv} = a_{vv} + f^2_{uv} + f_u(f_v f_{uv} - a_u),$$

$$(3.3)\qquad f_{uvvu} = -a_{uu} + f^2_{uv} + f_v(a_v + f_u f_{uv}).$$

whence $h^2 h_1$ has the two expressions

$$\begin{aligned} h^2 h_1 &= -a_u a_v + f_{uv}(f_v a_v - a_{vv}),\\ &= -a_u a_v + f_{uv}(a_{uu} - f_u a_u). \end{aligned}$$

Using the first form

$$(h^2 h_1)_u = f_u h^2 h_1.$$

Similarly, the second gives

$$(h^2 h_1)_v = f_v h^2 h_1,$$

whence

$$h^2 h_1 = k e^f,$$

with k an arbitrary constant. Since $k \neq 0$—unless $hh_1 = 0$, a special case we return to later—a transformation (T) can be found that will reduce k to unity and not violate the conditions (2.7).

From the first and last of (2.10) we find

$$(3.4)\qquad a = f_{uuuv}/f_{uv} - f_u f_{uuv}/f_{uv} - f_{uu}.$$

Eliminating a from the first two

$$(3.5)\qquad [f_{uu} + f_{vv} - (f_u^2 + f_v^2)/2]_{uv} = 0.$$

The conditions (3.4), (3.5) and

$$(3.6)\qquad h^2 h_1 = e^f$$

are equivalent to (2.10).

4. *Integration of* (Y). We turn now to the system (Y) which becomes

$$(4.1)\qquad \begin{aligned} y_{uu} &= y_{vv} + f_u y_u - f_v y_v + ay,\\ y_{uv} &= f_{uv} y \end{aligned}$$

and proceed to obtain the general integral of $y_{uv} = f_{uv} y$. This equation has equal invariants

$$h = k = f_{uv}. \tag{4.2}$$

Its first Laplace transform has the invariants

$$h_1 = f_{uv} - (\partial^2/\partial u \partial v) \log f_{uv}, \quad k_1 = h, \tag{4.3}$$

while, for the second transform,

$$h_2 = (f - \log h^2 h_1)_{uv}, \quad k_2 = h_1. \tag{4.4}$$

A glance at (3.6) shows that $h_2 = 0$, so that according to the general theory,* the general solution of $y_{uv} = f_{uv} y$ is

$$y = -aU + f_u U' + U'' + aV + f_v V' + V'', \tag{4.5}$$

in which U and V are arbitrary functions of u alone and v alone, respectively. The primes denote differentiation.

This value of y must also satisfy the first equation in (4.1). Direct substitution gives

$$\begin{aligned} & U^{\mathrm{IV}} + (2f_{uu} - f_u^2 - 2a) U'' \\ & \quad + (f_{uuu} - f_u f_{uu} - a_u) U' + (-a_{uu} + a_{vv} + a_u f_u - a_v f_v + a^2) U \\ \equiv\ & V^{\mathrm{IV}} + (2f_{vv} - f_v^2 + 2a) V'' \\ & \quad + (f_{vvv} - f_v f_{vv} + a_v) V' + (a_{vv} - a_{uu} - a_v f_v + a_u f_u + a^2) V. \end{aligned} \tag{4.6}$$

The coefficients of U'' and V'' are functions are functions of u alone and v alone, respectively, as (2.10) shows. Define

$$2U_0 \equiv 2f_{uu} - f_u^2 - 2a, \quad 2V_0 \equiv 2f_{vv} - f_v^2 + 2a. \tag{4.7}$$

Then equate the two expressions (3.2), (3.3). We find

$$a_{uu} + a_{vv} - f_u a_u - f_v a_v = 0 \tag{4.8}$$

so that the coefficient of U—and also of V—may be written indifferently

$$-2a_{uu} + 2a_u f_u + a^2, \quad 2a_{vv} - 2a_v f_v + a^2.$$

Differentiate the first form with respect to v, whence (2.10) shows the result to be zero. Similarly, the derivative of the second form with respect to u is zero. Thus the coefficient of U is a constant k and the condition (4.6) becomes

$$U^{\mathrm{IV}} + 2U_0 U'' + U'_0 U' + kU \equiv V^{\mathrm{IV}} + 2V_0 V'' + V'_0 V' + kV. \tag{4.9}$$

---

* Darboux, *Théorie Générale des Surfaces*, t. 2, Chap. 2.

Differentiating we find

$$(4.10)\qquad U^{V}+2U_0U'''+3U'_0U''+(U''_0+k)U'=0,$$
$$(4.11)\qquad V^{V}+2V_0V'''+3V'_0V''+(V''_0+k)V'=0,$$

both of which are self-adjoint ordinary differential equations.

We summarize our results: *The homogeneous point coördinates of the most general, non-degenerate, non-developable surface S_y referred to a conjugate net of plane curves with equal invariants are given by*

$$y^{(i)}=-aU_i+f_uU'_i+U''_i+aV_i+f_vV'_i+V''_i,\qquad (i=1,2,3,4),$$

where a and f are solutions of (2.10) *and U_i, V_i satisfy* (4.9).

5. *Integration of* (4.10). The equation (4.10) is of odd order and is also self-adjoint. Darboux has shown * that such an equation can always be expressed in the form

$$(5.1)\qquad f[(d/du)g(U)]=0,$$

where, in the present instance,

$$(5.2)\qquad \begin{aligned} f&\equiv\lambda U+\lambda_1U'+\lambda_2U'',\\ g&\equiv\mu U+\mu_1U'+\mu_2U'',\end{aligned}$$

and are adjoint expressions, that is,

$$(5.3)\qquad \mu=\lambda-\lambda'_1+\lambda_2'',\quad \mu_1=-\lambda_1+2\lambda'_2,\quad \mu_2=\lambda_2.$$

We desire to identify the equations (5.1) and (4.10). Comparison of coefficients of U^{V} indicates that we may take $\lambda_2=\mu_2=1$. Whence from (5.3), $\mu_1=-\lambda_1$ and $\mu=\lambda-\lambda'_1$. Direct calculation gives

$$(5.4)\quad \begin{aligned} f[(d/du)g(U)]\equiv{}& U^{V}+(2\lambda-4\lambda'_1-\lambda_1^2)U'''\\ &+3(\lambda'-2\lambda_1''-\lambda_1\lambda'_1)U''\\ &+(3\lambda''-4\lambda_1'''-3\lambda_1\lambda_1''+2\lambda_1\lambda'+\lambda^2-2\lambda\lambda'_1)U'\\ &+[\lambda'''-\lambda_1^{IV}+\lambda_1(\lambda''-\lambda_1''')+\lambda(\lambda'-\lambda_1'')]U.\end{aligned}$$

From the coefficients of U''', U'', U',

$$(5.5)\quad U_0=\lambda-2\lambda'_1-\lambda_1^2/2,\ k=(\lambda-\lambda'_1)^2+2(\lambda-\lambda'_1)''+2\lambda_1(\lambda-\lambda'_1)'.$$

Replace U_0 by

$$(5.6)\qquad U_0=k^{1/2}-(d^2/du^2)\log U_1-\tfrac{1}{2}(U'_1/U_1)^2$$

* *Théorie des Surfaces*, t. 2, p. 134.

where U_1 is a new function of u. Finally, choose:

$$\lambda_1 = U'_1/U_1, \quad \lambda = k^{1/2} + d^2 \log U_1/du^2, \tag{5.7}$$

whence

$$\begin{aligned} f(U) &\equiv (k^{1/2} + d^2 \log U_1/du^2)U + (U'_1/U_1)U' + U''; \\ g(U) &\equiv k^{1/2}U - (U'_1/U_1)U' + U''. \end{aligned} \tag{5.2'}$$

Darboux gives another form that is more useful in the sequel. We write

$$\begin{aligned} f(U) &\equiv (1/\alpha)(d/du)(1/\beta)(d/du)\alpha\beta \cdot U, \\ g(U) &\equiv \alpha\beta(d/du)(1/\beta)(d/du)(U/\alpha). \end{aligned} \tag{5.8}$$

Differentiating out and identifying with (5.2′) we find

$$\beta'/\beta = U'_1/U_1 - 2\alpha'/\alpha, \quad \alpha'' - (U'_1/U_1)\alpha' + k^{1/2}\alpha = 0.$$

Therefore the self-adjoint equation (4.10) may be written in the form (5.1) with f and g given by (5.8) and

$$\beta = U_1/\alpha^2; \tag{5.9}$$

$$\alpha'' - (U'_1/U_1)\alpha' + k^{1/2}\alpha = 0. \tag{5.10}$$

We may write down immediately five independent integrals of (4.10). They are

$$\begin{aligned} &u_1 = \alpha, \quad u_2 = \alpha\int \beta du, \quad u_3 = \alpha\int \beta du \int \frac{du}{\alpha\beta}, \\ &u_4 = \alpha\int \beta du \int \frac{du}{\alpha\beta}\int \frac{du}{\alpha\beta}, \quad u_5 = \alpha\int \beta du \int \frac{du}{\alpha\beta}\int \frac{du}{\alpha\beta}\int \beta du, \end{aligned} \tag{5.11}$$

where each integral includes those that follow. The integrals are to be regarded as having a common fixed lower limit u_0 and a common variable upper limit u.

In a similar fashion we write five independent integrals of (4.11).

$$\begin{aligned} &v_1 = \gamma, \quad v_2 = \gamma\int \delta\, dv, \quad v_3 = \gamma\int \delta\, dv \int \frac{dv}{\gamma\delta}, \\ &v_4 = \gamma\int \delta\, dv \int \frac{dv}{\gamma\delta}\int \frac{dv}{\gamma\delta}, \quad v_5 = \gamma\int \delta\, dv \int \frac{dv}{\gamma\delta}\int \frac{dv}{\gamma\delta}\int \delta\, dv, \end{aligned} \tag{5.12}$$

where

$$\delta = V_1/\gamma^2, \quad V_0 \equiv k^{1/2} - d^2 \log V_1/dv^2 - \tfrac{1}{2}(V'_1/V_1)^2, \tag{5.13}$$

$$\gamma'' - (V'_1/V_1)\gamma' + k^{1/2}\gamma = 0, \tag{5.14}$$

with corresponding conventions with respect to the integrals.

A pair u_i, v_j, when substituted in (4.5), will give a solution of (Y) provided

$$(5.15) \qquad F(u_i) \equiv G(v_j),$$

in which $F(U)$ is a notation for the left member of (4.9) and $G(V)$ for the right member. We verify that

$$(5.16) \quad \begin{aligned} &F(u_1)=0,\ F(u_2)=0,\ F(u_3)=k^{1/2},\ F(u_4)=\alpha'/U_1+k^{1/2}\int_{u_0}^{u}\frac{\alpha}{U_1}\,du,\\ &F(u_5)=\frac{U_1}{\alpha'}\int_{u_0}^{u}\frac{U_1}{\alpha^2}\,du+k^{1/2}\int_{u_0}^{u}\frac{\alpha}{U_1}\,du\int_{u_0}^{u}\frac{U_1}{\alpha^2}\,du+1/\alpha,\end{aligned}$$

with similar results for $G(v_j)$. We need merely to replace α, U_1, u by γ, V_1, v. $F(u_4)$ and $F(u_5)$ are constants for their derivatives are zero. Hence

$$F(u_4)=\alpha'(u_0)/U_1(u_0), \quad F(u_5)=1/\alpha(u_0).$$

In order that

$$(5.17) \qquad F(u_i)=G(v_i) \qquad (i=1,2,3,4,5),$$

we agree that $\alpha'(u_0)=U_1(u_0)$, $\alpha(u_0)=\gamma(v_0)=1$, $\gamma'(v_0)=V_1(v_0)$. Hence (5.10) may be written in the equivalent form

$$\alpha'/U_1+k^{1/2}\int_{u_0}^{u}(\alpha/U_1)\,du=1.$$

These five pairs (u_i, v_i) give five solutions $y^{(i)}$ which, of course, cannot be independent. We select four pairs so as to obtain four linearly independent $y^{(i)}$ from (4.5). The following may be used

$$(5.18) \quad \begin{aligned} &\bar u_1=2u_1, && \bar u_2=2u_3, && \bar u_3=u_4, && \bar u_4=2u_5,\\ &\bar v_1=v_1, && \bar v_2=v_3+k^{1/2}v_4, && \bar v_3=v_4, && \bar v_4=v_4+v_5.\end{aligned}$$

In fact, $F(\bar u_i)=G(\bar v_i)$, $(i=1,2,3,4)$ and the determinant

$$|\,y_{vv}^{(i)}, y_u^{(i)}, y_v^{(i)}, y^{(i)}\,|$$

is not identically zero. This latter fact may be verified most conveniently by using the particular solution

$$(5.19) \qquad a=-4/(u^2-v^2), \quad f=\log 2/(u^2-v^2)^2.$$

with which we may use

$$(5.20) \qquad u_0=0, \quad U_1=1,$$

whence $\alpha=u+1$. In this case the above determinant has the value $8/(u^2-v^2)^2$.

The closing phrase in the theorem at the end of paragraph 4 should now be amended to read "and $U_i = \bar{u}_i$, $V_i = \bar{v}_i$" where $\bar{u}_i$ and $\bar{v}_i$ are given by the expressions (5.18), (5.11), and (5.12).

6. *Exceptional case* $h_1 = 0$, $h \neq 0$. The Liouville equation $h_1 = 0$ has for its general solution

$$f = \log 2U'_0V'_0/U_1V_1(U_0 + V_0)^2, \tag{6.1}$$

where the U's and V's are arbitrary functions of u alone and v alone, respectively. From the integrability conditions (2.10) the value of a_{uv} may be computed from either of the first two conditions. Then the third gives

$$a = \partial^2 \log U_1/\partial u^2 + (U'_1/U_1)(\partial/\partial u) \log 2U'_0V'_0/(U_0 + V_0)^2,$$

or

$$a = -\partial^2 \log V_1/\partial v^2 - (V'_1/V_1)(\partial/\partial v) \log 2U'_0V'_0/(U_0 + V_0)^2,$$

whence, equating,

$$[U'_1(2U'_0V'_0)/U_1(U_0 + V_0)^2]_u = -[V'_1(2U'_0V'_0)/V_1(U_0 + V_0)^2]_v. \tag{6.2}$$

Hence there exists a function $\phi(u, v)$ such that

$$\phi_v = U'_1(2U'_0V'_0)/U_1(U_0 + V_0)^2, \quad \phi_u = -V'_1(2U'_0V'_0)/V_1(U_0 + V_0)^2.$$

Integrating the first and substituting the result in the second, we obtain

$$\begin{aligned} d^2 \log U_1/du^2 + (U'_1/U_1)[U''_0/U'_0 - U'_0/(U_0 + V_0)] \\ - U'_2(U_0 + V_0)/2U'_0 - V'_1V'_0/V_1(U_0 + V_0) = 0. \end{aligned} \tag{6.3}$$

In the same way the integration of the second gives

$$\begin{aligned} d^2 \log V_1/dv^2 + (V'_1/V_1)[V''_0/V'_0 - V'_0/(U_0 + V_0)] \\ + V'_2(U_0 + V_0)/2V'_0 - U'_1U'_0/U_1(U_0 + V_0) = 0. \end{aligned} \tag{6.4}$$

In these equations U_2, V_2 are arbitrary functions entering through the integrations. From (6.2), (6.3), (6.4)

$$U'_2/U'_0 = V'_2/V'_0 = \text{constant} = k,$$

whence

$$U_2 = kU_0 + k_3, \quad V_2 = kV_0 + 2k_1 + k_3,$$

and

$$\phi = -U'_1(2U'_0)/U_1(U_0 + V_0) + kV_0 + k_3,$$

or

$$\phi = V'_1(2V'_0)/V_1(U_0 + V_0) + kV_0 + 2k_1 + k_3.$$

By equating these two expressions for ϕ we find

$$(6.5)\qquad 2U'_0U'_1/U_1 - kU_0^2 + 2k_1U_0 = -2V'_0V'_1/V_1 - kV_0^2 - 2k_1V_0 = \text{constant} = -k_2.$$

Differentiating with respect to u and dividing by $2U'_0$

$$d^2 \log U_1/du^2 + U'_1U_0''/U_1U'_0 - kU_0 + k_1 = 0.$$

In a similar manner

$$d^2 \log V_1/dv^2 + V'_1V_0''/V_1V'_0 + kV_0 + k_1 = 0.$$

The value of a may now be written, using the last three equations,

$$(6.6)\qquad (U_0 + V_0)a = kU_0V_0 + k_1(U_0 - V_0) + k_2.$$

We have thus obtained the formulas for the functions a and f. They are given by (6.6) and (6.1). The values of U_1, V_1 which enter therein are found from (6.5). The arbitraries thus consist of U_0, V_0 and five constants.

Let us now investigate the problem of integrating the system (Y). Since the Laplace invariant h_1 of the second equation is zero, the general integral is

$$(6.7)\qquad y = (h_u/h)U + U' + (h_v/h)V + V',$$

in which U, V are arbitrary functions of u alone and v alone. Direct substitution in the first equation gives

$$(6.8)\qquad \begin{aligned} &U''' + (U'_1/U_1)U'' + [2\{U_0, u\} - d^2 \log U_1/du^2]\,U' \\ &\quad + [\{U_0, u\}' + 2(U'_1/U_1)\{U_0, u\} + d^3 \log U_1/du^3]\,U \\ &= V''' + (V'_1/V_1)V'' + [2\{V_0, v\} - d^2 \log V_1/dv^2]\,V' \\ &\quad + [\{V_0, v\}' + 2(V'_1/V_1)\{V_0, v\} + d^3 \log V_1/dv^3]\,V, \end{aligned}$$

where we use the familiar notation $\{U_0, u\}$ for the Schwarzian derivative of U_0 with respect to u. Differentiating with respect to u

$$(6.9)\qquad \begin{aligned} &U^{\mathrm{IV}} + (U'_1/U_1)U''' + 2\{U_0, u\}U'' + [3\{U_0, u\}' + 2(U'_1/U_1)\{U_0, u\}]\,U' \\ &\quad + [\{U_0, u\}'' + (U'_1/U_1)\{U_0, u\}']\,U = 0, \end{aligned}$$

since

$$(U'_1/U_1)\{U_0, u\}' + 2(d^2 \log U_1/du^2)\{U_0, u\} + d^4 \log U_1/du^4 = 0.$$

There is a similar equation for the function V.

Although the equation (6.9) is not self-adjoint it may be written in the form

$$(6.9')\qquad (1/U_1)(d/du)U'_0U_1(d/du)(1/U'_0)(d/du)(1/U'_0)(d/du)U'_0U = 0.$$

Hence four linearly independent solutions are

$$(6.10)\quad u_1 = 1/U'_0, \quad u_2 = U_0/U'_0, \quad u_3 = U_0^2/2U'_0,$$
$$u_4 = u_1 \int_{u_0}^{u} (u_3/U_1)\,du - u_2 \int_{u_0}^{u} (u_2/U_1)\,du + u_3 \int_{u_0}^{u} (u_1/U_1)\,du.$$

Similarly, we may use

$$(6.11)\quad v_1 = 1/V'_0, \quad v_2 = -V_0/V'_0, \quad v_3 = V_0^2/2V'_0,$$
$$v_4 = v_1 \int_{v_0}^{v} (v_3/V_1)\,dv - v_2 \int_{v_0}^{v} (v_2/V_1)\,dv + v_3 \int_{v_0}^{v} (v_1/V_1)\,dv,$$

as a set of linearly independent solutions of the V-equation corresponding to (6.9).

Call the left member of (6.8) $F(U)$ and the right member $G(V)$. Then

$$(6.12)\quad \begin{array}{llll} F(u_1) = k, & F(u_2) = k_1, & F(u_3) = -k_2/2, & F(u_4) = 1/U_1(u_0), \\ G(v_1) = -k, & G(v_2) = -k_1, & G(v_3) = k_2/2, & G(v_4) = 1/V_1(v_0), \end{array}$$

the k's entering through (6.5).

In order to obtain four linearly independent $y^{(i)}$ that will be solutions of the equations (4.1) we select pairs of linear expressions $\bar{u}_i$, $\bar{v}_i$ of u_i and v_i, respectively, such that $F(\bar{u}_i) = G(\bar{v}_i)$ while the determinant

$$|\, y_{uv}^{(i)},\ y_u^{(i)},\ y_v^{(i)},\ y^{(i)} \,| \neq 0,$$

the $y^{(i)}$ being computed from (6.7). To accomplish this we take

$$(6.13)\quad \begin{array}{ll} \bar{u}_1 = u_1 + a_1 u_2, & \bar{v}_1 = v_1 + \alpha_1 v_2, \\ \bar{u}_2 = u_2 + a_2 u_3, & \bar{v}_2 = v_2 + \alpha_2 v_3, \\ \bar{u}_3 = u_3 + a_3 u_4, & \bar{v}_3 = v_3 + \alpha_3 v_4, \\ \bar{u}_4 = u_4 + a_4 u_3, & \bar{v}_4 = v_4 + \alpha_4 v_3, \end{array}$$

where

$$(6.14)\quad a_1 + \alpha_1 = -2k/k_1, \; a_2 + \alpha_2 = 4k_1/k_2, \; a_3 = \alpha_3 + k_2 \neq 0, \; a_4 + \alpha_4 = 0.$$

Hence

$$(6.15)\quad \begin{aligned} y^{(1)} &= -4/(U_0 + V_0) + 2k(U_0 - V_0)/k_1(U_0 + V_0), \\ y^{(2)} &= -2(U_0 - V_0)/(U_0 + V_0) + 4k_1 U_0 V_0/k_2(U_0 + V_0), \\ y^{(3)} &= 2U_0V_0/(U_0 + V_0) + k_2\left(\frac{2}{U_0 + V_0}\int_{v_0}^{v}\frac{v_3}{V_1}\,dv - \frac{U_0 - V_0}{U_0 + V_0}\right. \\ &\qquad \left.\int_{v_0}^{v}\frac{v_2}{V_1}\,dv - \frac{U_0V_0}{U_0 + V_0}\int_{v_0}^{v}\frac{v_1}{V_1}\,dv\right) \\ y^{(4)} &= -\frac{2}{U_0 + V_0}\left(\int_{u_0}^{u}\frac{u_3}{U_1}\,du + \int_{v_0}^{v}\frac{v_3}{V_1}\,dv\right) + \frac{U_0 - V_0}{U_0 + V_0}\left(\int_{u_0}^{u}\frac{u_2}{U_1}\,du\right. \\ &\qquad \left. + \int_{v_0}^{v}\frac{v_2}{V_1}\,dv\right) + \frac{U_0V_0}{U_0 + V_0}\left(\int_{u_0}^{u}\frac{u_1}{U_1}\,du + \int_{v_0}^{v}\frac{v_1}{V_1}\,dv\right). \end{aligned}$$

The above determinant is not identically zero for it does not vanish when $a = 0$.

We turn now to some geometrical considerations. From (1.2)

$$(6.16)\qquad \epsilon = y_{uu} - (h_u/h)y_u, \quad \sigma = z_v - (h_v/h)z,$$

so that

$$(6.17)\qquad \sigma = \epsilon - ay - (f_u - h_u/h)y_u + (f_v - h_v/h)z.$$

By direct computation

$$(6.18)\qquad \begin{aligned} \sigma_u &= 0, \quad \sigma_v = (f_v - h_v/h)\sigma, \\ \epsilon_v &= 0, \quad \epsilon_u = (f_u - h_u/h)\epsilon, \end{aligned}$$

from which we see that the developables S_z and S_ρ reduce to cones whose vertices are the fixed points P_σ and P_ϵ, respectively. Ordinarily these vertices do not coincide. The relation (6.17) shows this to be the case only if

$$a = 0, \quad f_u - h_u/h = 0, \quad f_v - h_v/h = 0$$

which conditions are equivalent to the single one

$$(6.19)\qquad a = 0.$$

The surfaces for which $a = 0$ have been studied elsewhere.* We note that the formulas (6.15) are still valid. We merely allow k, k_1, k_2 to approach zero in such a way that k/k_1 and k_1/k_2 approach zero. In this case the cones reduce to coincident quadric cones. In fact,

$$\begin{aligned} z^{(1)} &= 4V'_0/(U_0 + V_0)^2, \quad z^{(2)} = 4U_0V'_0/(U_0 + V_0)^2, \\ z^{(3)} &= 2U_0^2V'_0/(U_0 + V_0)^2, \end{aligned}$$

while

$$\begin{aligned} \rho^{(1)} &= 4U'_0/(U_0 + V_0)^2, \quad \rho^{(2)} = -4U'_0V_0/(U_0 + V_0)^2, \\ \rho^{(3)} &= 2U'_0V_0^2/(U_0 + V_0)^2, \end{aligned}$$

whence both $z^{(4)}$ and $\rho^{(4)}$ satisfy

$$(6.20)\qquad 2x_1x_3 - x_2^2 = 0,$$

which is a quadric cone with vertex at $(0, 0, 0, 1)$.

It should be noted that these coördinates refer to an arbitrary tetrahedron of reference. To find the equation of the cone referred to the local tetra-

---

* Lane, "Conjugate Systems with Indeterminate Axis Curves," *American Journal of Mathematics*, Vol. 43 (1921), p. 52. The totality of surfaces possessing conjugate systems with indeterminate axis curves forms, to be sure, a larger class than those we are considering here.

hedron whose vertices are P_y, P_z, P_ρ, P_σ we express the derivatives of z in terms of y, z, ρ, σ. Those which we need are

$$z_u = f_{uv}y, \qquad z_{uu} = f_{uuv}y + f_{uv}\rho,$$
$$z_{uv} = f_{uvv}y + f_{uv}z, \quad z_{vv} = (f_{vv} + f_v^2)z + f_v\sigma.$$

Placing these values in the expansion

$$\begin{aligned}(6.21) \qquad Z &= z + z_u \cdot \Delta u + z_v \cdot \Delta v \\ &\quad + \tfrac{1}{2}(z_{uu} \cdot \Delta u^2 + 2z_{uv} \cdot \Delta u \Delta v + z_{vv} \cdot \Delta v^2) + \cdots \\ &= z_1 y + z_2 z + z_3 \rho + z_3 \sigma\end{aligned}$$

we have

$$\begin{aligned} & z_1 = f_{uv} \cdot \Delta u + (f_{uuv}/2)\Delta u^2 + f_{uvv} \cdot \Delta u \Delta v + \cdots, \\ (6.22) \qquad & z_2 = 1 + \cdots, \\ & z_3 = (f_{uv}/2)\Delta u^2 + \cdots. \end{aligned}$$

In each of the above expressions the omitted terms are of higher order than those given.

Since the vertex of the cone is at $(0, 0, 0, 1)$ and the tangent planes at $(0, 1, 0, 0)$ and $(0, 0, 1, 0)$ are $x_3 = 0$ and $x_2 = 0$, respectively, the equation must be of the form

$$(6.23) \qquad y_1^2 - 2\lambda y_2 y_3 = 0.$$

To determine λ we insert the values of z in (6.23) and demand that all terms of the second order in Δu and Δv disappear. Hence

$$\lambda = f_{uv} = 2U'_0V'_0/(U_0 + V_0)^2.$$

7. *Exceptional Case* $h = f_{uv} = 0$. When $f_{uv} = 0$, the integrability conditions (I) give

$$(7.1) \qquad f = U_0 + V_0, \qquad a = \text{constant} = \alpha.$$

The general solution of the second equation of (Y) is

$$(7.2) \qquad y = U + V,$$

whence substitution in the first gives

$$(7.3) \qquad U'' - U'_0U' - \alpha U \equiv V'' - V'_0V' + \alpha V = \textbf{constant}.$$

For a moment we turn to geometrical considerations. The fundamental equations (Δ) show that

$$z_u = 0, \quad z_{vv} = (f_{vv} - \alpha)z + f_v z_v,$$

so that the developable S_z becomes a straight line. Similarly, for S_ρ. These two lines do not meet in general. In fact, any point on the line L_z is given by

$$(7.4) \qquad z_v + \kappa z = y_{vv} + \kappa y_v,$$

while an arbitrary point of L_ρ is of the form

$$(7.5) \qquad \rho_u + \lambda\rho = y_{uu} + \lambda y_u = y_{vv} + (\lambda + U'_0) y_u - V'_0 y_v + \alpha y.$$

For intersection all second order determinants of the matrix

$$\begin{pmatrix} 1 & 0 & \kappa & 0 \\ 1 & \lambda + U'_0 & -V'_0 & \alpha \end{pmatrix}$$

must vanish. Consequently,

$$\alpha = 0, \quad \lambda = -U'_0, \quad \kappa = -V'_0,$$

and we may say that *a necessary and sufficient condition for the lines L_z, L_ρ to meet is that $a = \alpha = 0$.*

Now for the integration of the equation (7.3). Differentiation with respect to u yields

$$(7.6) \qquad U''' - U'_0 U'' - (U_0'' + \alpha) U' = 0.$$

Replace the arbitrary function U_0 by

$$(7.7) \qquad U'_0 = U_1''/U'_1 - \alpha U_1/U'_1,$$

so that $U = U_1$ is a particular integral of (7.6). Then make the usual transformation $U' = U'_1 \bar{U}$ and integrate. In this way we find three independent integrals of (7.6) to be

$$(7.8) \quad u_1 = 1, \quad u_2 = U_1, \quad u_3 = U_1 \int_{u_0}^{u} (U_2/U'_1)\, du - \int_{u_0}^{u} (U_1 U_2/U'_1)\, du,$$

where

$$(7.9) \qquad U_2 = \exp\left[-\alpha \int_{u_0}^{u} (U_1/U'_1)\, du\right].$$

In the same manner, three independent integrals of the equation in V, corresponding to (7.6), are

$$(7.10) \quad v_1 = 1, \quad v_2 = V_1, \quad v_3 = V_1 \int_{v_0}^{v} (V_2/V'_1)\, dv - \int_{v_0}^{v} (V_1 V_2/V'_1)\, dv,$$

in which

$$(7.11) \qquad V_0 = V_1''/V_1 + \alpha V_1/V'_1, V_2 = \exp\left[\alpha \int_{u_0}^{v} (V_1/V'_1)\, dv\right].$$

Now if the left member of (7.3) be denoted by $F(U)$

$$F(u_1) = -\alpha, \quad F(u_2) = 0, \quad F(u_3) = U_2 - \int_{u_0}^{u} U'_2 du = U_2(u_0).$$

We may suppose $U_2(u_0) = 1$. Similarly,

$$G(v_1) = \alpha, \quad G(v_2) = 0, \quad G(v_3) = 1.$$

The following combinations of u_i and v_i give the four solutions $y^{(i)}$

$$(7.12) \quad \begin{aligned} y^{(1)} &= u_1 + u_2 + \alpha u_3 + v_1 - \alpha v_3 = 2 + U_1 + \alpha(u_3 - v_3), \\ y^{(2)} &= u_1 + \alpha u_3 + v_1 + v_2 - \alpha v_3 = 2 + V_1 + \alpha(u_3 - v_3), \\ y^{(3)} &= u_2 + v_2 = U_1 + V_1, \\ y^{(4)} &= u_3 + v_3 = U_1 \int_{u_0}^{u} \frac{U_2}{U'_1} du - \int_{u_0}^{u} \frac{U_1 U_2}{U'_1} du \\ &\qquad + V_1 \int_{v_0}^{v} \frac{V_2}{V'_1} dv - \int_{v_0}^{v} \frac{V_1 V_2}{V'_1} dv, \end{aligned}$$

which are linearly independent as may be verified by computing the value of $| y_{vv}^{(i)}, y_u^{(i)}, y_v^{(i)}, y^{(i)} |$ when $\alpha = 0$.

The finite equations of the lines L_s and L_ρ may be found by substituting the expressions (7.12) in (1.2) and eliminating the functions U_1, V_1. The results are, respectively,

$$(7.13) \quad x_1 + \alpha x_4 = x_1 - x_2 + x_3 = 0, \quad x_2 - \alpha x_4 = x_1 - x_2 - x_3 = 0.$$

If $\alpha = 0$ the lines meet in the point $(0, 0, 0, 1)$. In the contrary case there is no point of intersection.

A TERNARY ANALOGUE OF ABELIAN GROUPS.

By D. H. Lehmer.*

Introduction. A class K with an operation called multiplication applied to *pairs* of elements of K is an abstract group provided certain postulates are satisfied. Unfortunately the name group does not suggest the *binary* character of multiplication. The entities with which this note is concerned are similar to groups, the class K being subjected to a *ternary* operation however. For want of a more descriptive name we have called them *triplexes*. The need for their consideration arose in an attempt to obtain solutions of a pair of functional equations, but their investigation would seem justified by intrinsic interest especially when compared with Abelian groups.

The system of postulates on which we base our investigation is modelled after Hurwitz's † system for Abelian groups, and accordingly differs somewhat from the definitions of a group found in most treatises on group theory. In this way we reduce the proofs of several theorems to a minimum and have at the same time a more perfect system from a strictly logical point of view. The role that Abelian groups play in this theory is described in § 3. The rest of the paper deals with finite triplexes and concepts analogous to the fundamental notions of the theory of Abelian groups such as order and inverse of an element, sub-group, cyclic group, quotient group, etc. Two notions however are conspicuous by their absence, the unit and the basis. Other facts such as the existence of triplexes with no subtriplex stand out as being different from what one might expect from the theory of Abelian groups.

1. *Definition of a Triplex.* A class K with an operation $\cdot$ between triplets of elements is called a triplex if the following postulates hold.

Postulate I.
$$\begin{aligned}(a\cdot b\cdot c)\cdot d\cdot e &= d\cdot(a\cdot b\cdot c)\cdot e = d\cdot e\cdot(a\cdot b\cdot c)\\ &= (a\cdot b\cdot d)\cdot c\cdot e = (a\cdot b\cdot e)\cdot c\cdot d = (a\cdot c\cdot d)\cdot b\cdot e\\ &= (a\cdot c\cdot e)\cdot b\cdot d = (a\cdot d\cdot e)\cdot b\cdot c = (b\cdot c\cdot d)\cdot a\cdot e\\ &= (b\cdot c\cdot e)\cdot a\cdot d = (b\cdot d\cdot e)\cdot a\cdot c = (c\cdot d\cdot e)\cdot a\cdot b\end{aligned}$$

provided a, b, c, d, e, and all the expressions in which these letters are involved in the above, belong to K.

Postulate II. If a, b, c, belong to K, there is an element x of K such that $a\cdot b\cdot x = c$.

* National Research Fellow.

† *Annals of Mathematics*, Ser. 2, Vol. 15, p. 93.

The number of elements in K is called the order of the triplex and is specified when necessary by adding one of the postulates:

Postulate III_1. K contains n elements.

Postulate III_2. K contains infinitely many elements.

According as III_1 or III_2 holds, the triplex is called finite or infinite. We proceed to deduce 5 fundamental theorems from Postulates I and II.

THEOREM 1. *If a_1, a_2, a_3, and one of the six expressions obtained from $a_1 \cdot a_2 \cdot a_3$ by permuting the letters, belong to K, then these six expressions are equal.*

Proof. Suppose for definiteness that $a_1 \cdot a_2 \cdot a_3$ belongs to K. Choose x_1, so that $a_2 \cdot a_3 \cdot x_1 = a_1$. In Postulate I set $a = d = a_2$, $b = e = a_3$, $c = x_1$. The first two equalities of Postulate I become

$$a_1 \cdot a_2 \cdot a_3 = a_2 \cdot a_1 \cdot a_3 = a_2 \cdot a_3 \cdot a_1.$$

Next choose x_2 so that $a_1 \cdot a_3 \cdot x_2 = a_2$. In Postulate I set $a = e = a_1$, $b = d = a_3$, $c = x_2$. Then we have, since $a_2 \cdot a_3 \cdot a_1$ is now in K

$$a_2 \cdot a_3 \cdot a_1 = a_3 \cdot a_2 \cdot a_1 = a_3 \cdot a_1 \cdot a_2.$$

Finally choose x_3 so that $a_1 \cdot a_2 \cdot x_3 = a_3$ and set $a = d = a_1$, $b = e = a_2$, $c = x_3$. Then since $a_3 \cdot a_1 \cdot a_2$ belongs to K, we have

$$a_3 \cdot a_1 \cdot a_2 = a_1 \cdot a_3 \cdot a_2 = a_1 \cdot a_2 \cdot a_3.$$

Hence the theorem.

Definition I. $a \cdot b \cdot c \cdot d \cdot e = (a \cdot b \cdot c) \cdot d \cdot e.$

THEOREM 2. *If the letters a, b, c, d, e, and their combinations appearing in Postulate I belong to K, then $a \cdot b \cdot c \cdot d \cdot e$ is unaltered by permuting the letters.*

Proof. Consider the 10 expressions obtained by striking out the expressions $d \cdot (a \cdot b \cdot c) \cdot e$ and $d \cdot e \cdot (a \cdot b \cdot c)$ from the list of Postulate II. They are all of the form

$$(a_1 \cdot a_2 \cdot a_3) \cdot a_4 \cdot a_5 = a_1 \cdot a_2 \cdot a_3 \cdot a_4 \cdot a_5.$$

In each of these 10 expressions we may, by Theorem 1, permute the letters in the parenthesis and interchange the two letters outside the parenthesis. We obtain 120 equal expressions no two of which involve the 5 letters exactly in the same order. Applying Definition I, we obtain all the $5!$ permutations of $a \cdot b \cdot c \cdot d \cdot e$. These expressions are equal; hence the theorem.

THEOREM 3. *If a, b, c belong to K, then* $a \cdot b \cdot c$ belongs to K.

Proof. We may choose elements x, y, z, and w so that

$$a = b \cdot c \cdot x, \qquad x = a \cdot c \cdot y, \qquad a = y \cdot c \cdot z, \qquad z = a \cdot c \cdot w.$$

Then by Theorem 2

$$a = y \cdot c \cdot z = y \cdot c \cdot a \cdot c \cdot w = (a \cdot c \cdot y) \cdot w \cdot c = x \cdot w \cdot c.$$

But $z = a \cdot c \cdot w = b \cdot c \cdot x \cdot c \cdot w = (x \cdot w \cdot c) \cdot b \cdot c = a \cdot b \cdot c$. Thus $a \cdot b \cdot c$ is equal to an element z of K. Hence the theorem.

THEOREM 4. *The element x of Postulate II is unique.*

Proof. Suppose the theorem is false. It is then possible to find two distinct elements x_1 and x_2 of K so that

$$c = a \cdot b \cdot x_1 = a \cdot b \cdot x_2.$$

Let y be so chosen that $x_2 = a \cdot x_1 \cdot y$. Then

$$c = a \cdot b \cdot x_2 = a \cdot b \cdot a \cdot x_1 \cdot y = (a \cdot b \cdot x_1) \cdot a \cdot y = c \cdot a \cdot y.$$

Now let z be chosen so that $x_1 = a \cdot c \cdot z$. Then

$$x_1 = a \cdot z \cdot c = a \cdot z \cdot c \cdot a \cdot y = a \cdot x_1 \cdot y = x_2.$$

But this is impossible since x_1 and x_2 are distinct.

THEOREM 5. *If $a \cdot b \cdot x_1 = a \cdot b \cdot x_2$, then $x_1 = x_2$.*

Proof. This is in fact another statement of Theorem 4.

2. *Associated Elements.*

THEOREM 6. *If a and b are any elements of K, then the solution x of $a \cdot b \cdot x = b$ depends on a alone.*

Proof. Let c be an element of K. Then since $a \cdot b \cdot x = b$ we have

$$a \cdot (a \cdot b \cdot x) \cdot c = a \cdot b \cdot c = a \cdot b \cdot (a \cdot c \cdot x).$$

Hence by Theorem 5, $c = a \cdot c \cdot x$. Thus x does not depend on b.

We write $x = a'$ and call a' the element associated with a. By Theorem 4, a' is unique. Obviously $(a')' = a$. Also if a and a' both occur in an expression with $\cdot$ they may be cancelled.

THEOREM 7. $a' \cdot b' \cdot c' = (a \cdot b \cdot c)'$.

Proof. Let $a' \cdot b' \cdot c' = x_1$, and $(a \cdot b \cdot c)' = x_2$ and let d be any element

in K. Then $a \cdot b \cdot c \cdot d \cdot x_1 = d = (a \cdot b \cdot c) \cdot d \cdot x_2$. Hence by Theorem 5, $x_1 = x_2$, which is the theorem.

3. *Abelian Groups.* Given a class K and an operation $\cdot$, applied to pairs of elements under which K is an Abelian group, then it follows that K is also a triplex under the operation $\cdot$ applied to triplets of elements. In fact such a class satisfies Postulates I and II. In such a case the triplex is called the extension of the group. Obviously every Abelian group has only one extension. However some triplexes are the extensions of more than one Abelian group. For example we may take for K the class of all positive and negative rational numbers (0 excepted). The two groups under the two operations

$$a \cdot b = ab \qquad \text{and} \qquad a \cdot b = -ab$$

when extended give the same triplex since in both cases $a \cdot b \cdot c = abc$. It is possible to choose a class K and an operation $\cdot$ which when applied to pairs of elements does not give an Abelian group but nevertheless when $\cdot$ is applied to triplets we obtain a triplex. Consider for example the class K of all positive and negative odd integers. Under the operation $a \cdot b = a + b$, the class K is not a group since there is no closure property. Under the operation $a \cdot b \cdot c = a + b + c$ however, K is a triplex. Some triplexes are not the extension of any Abelian group. The simplest example of such a triplex is the following. Let the class K consist of two elements a and b. The operation $\cdot$ is defined by the following table.

$$\begin{array}{ll} a \cdot a \cdot a = b & a \cdot a \cdot b = a \cdot b \cdot a = b \cdot a \cdot a = a \\ b \cdot b \cdot b = a & a \cdot b \cdot b = b \cdot a \cdot b = b \cdot b \cdot a = b. \end{array}$$

Then K is a triplex under $\cdot$. But it is impossible to interpret $a \cdot a$, $a \cdot b$, and $b \cdot b$ in a consistent way. To give a concrete example of such a triplex set $a = i$, $b = -i$ and $a \cdot b \cdot c = abc$.

THEOREM 8. *If a triplex contains an element ϵ equal to its own associate ϵ', then the triplex is an extension of an Abelian group of which ϵ is the unit.*

Proof. For every pair (a, b) of elements of K there is a unique third element defined by $a \cdot b \cdot \epsilon = c$. We may express this fact by writing $a \cdot b = c$ and this notation is consistent because

$$(a \cdot b) \cdot d = (a \cdot b \cdot \epsilon) \cdot d \cdot \epsilon = a \cdot b \cdot d \cdot \epsilon \cdot \epsilon' = a \cdot b \cdot d.$$

Applying $\cdot$ to pairs of elements of K, we obtain an Abelian group of which as we have just seen, the triplex is an extension. Moreover

$$a \cdot \epsilon = a \cdot \epsilon \cdot \epsilon = a \cdot \epsilon \cdot \epsilon' = a.$$

Hence ϵ is the unit of the group. This proves the theorem.

THEOREM 9. *If a triplex is the extension of exactly r distinct Abelian groups, then the class K contains exactly r distinct elements $\epsilon_1, \epsilon_2, \cdots, \epsilon_r$ such that each is equal to its associate. In fact these elements are the units of the r groups.*

Proof. Let G_ν be the ν-th group $(\nu = 1, 2, \cdots, r)$ and let ϵ_ν be its unit. Let us symbolize its operation by $(ab)_\nu$. Then since $a \cdot b \cdot c = (a(bc)_\nu)_\nu$ we have for $b = c = \epsilon_\nu$

$$a \cdot \epsilon_\nu \cdot \epsilon_\nu = (a(\epsilon_\nu \epsilon_\nu)_\nu)_\nu = (a\epsilon_\nu)_\nu = a$$

But $a \cdot \epsilon_\nu \cdot \epsilon'_\nu = a$, hence by Theorem 5, $\epsilon_\nu = \epsilon'_\nu$. Suppose that these ϵ's are not all distinct, so that $\epsilon_\lambda = \epsilon_\mu$, then

$$(ab)_\lambda = a \cdot b \cdot \epsilon_\lambda = ((ab)_\lambda \epsilon_\lambda)_\lambda = ((ab)_\mu \epsilon_\lambda)_\mu = (ab)_\mu.$$

But this is impossible since a and b are quite arbitrary, and G_λ and G_μ are distinct. Hence the ϵ's are distinct. Since a group contains only one unit it follows from Theorem 8 that our triplex contains not more than r distinct elements each equal to its own associate.

4. *Finite Triplexes.* In what follows we suppose that Postulate III_1 holds so that the triplex is of a finite order n.

THEOREM 10. *If the order n of a triplex is odd then the triplex is an extension of an Abelian group.*

Proof. With every element a of K we can associate a'. Since n is odd at least one element is its own associate. Hence the theorem follows from Theorem 8.

5. *The Degree.* Consider any element a and its "powers."

$$a, \quad a \cdot a \cdot a = a^3, \quad a^3 \cdot a \cdot a = a^5, \cdots. \tag{1}$$

Since the order of the triplex is finite this sequence of powers of a contains a finite number of distinct elements. Let

$$a^{2k+1} = a^{2l+1} = a^{2m+1}, \qquad (k < l < m).$$

Then we have

$$a \cdot a \cdot a^{2k+1} = a^{2l+1} \cdot a^{2(l-k)+1} \cdot a = a \cdot a \cdot a^{2l+1}.$$

Hence by Theorem 5,

$$a = a^{2(l-k)+1} = a^{2(m-k)+1}.$$

Hence the sequence (1) contains two powers of a equal to a. If

$$a = a^{2p+1} = a^{2q+1},$$

then it follows in the same way that $a = a^{2(p\pm q)+1}$. Hence the set of all numbers r for which $a^{2r+1} = a$ is a module and coincides with all integer multiples of an integer α. Hence we have

THEOREM 11. *For each element a there exists an integer α such that $a^{2k+1} = a^{2l+1}$ if and only if $k \equiv l \pmod{\alpha}$.*

The number α is called the degree of a.

THEOREM 12. *$a' = a^{2\alpha-1}$ where α is the degree of a.*

Proof. Let b be an arbitrary element and consider the expression

$$a \cdot a \cdot (a \cdot b \cdot a^{2\alpha-1}) = a \cdot b \cdot a^{2\alpha+1} = a \cdot a \cdot b.$$

Hence $b = a \cdot b \cdot a^{2\alpha-1}$, and $a^{2\alpha-1}$ is the unique element associated with a. Moreover a and a' are of the same degree.

6. *Sub-Triplex.* Let T be a triplex defined by a class K and an operation $\cdot$. If there is a triplex T_1 defined by a sub-class of K and the same operation $\cdot$, T_1 is called a sub-triplex of T. For example, the elements $a, a^3, a^5, \cdots, a^{2\alpha-1}$ form, under $\cdot$, a sub-triplex, provided, of course, $\alpha < n$.

Let $u_1, u_2, u_3, \cdots, u_r$, be any sub-triplex T_1 of T. Consider u_m and the following sequences

$$a \cdot u_m \cdot u_1, \quad a \cdot u_m \cdot u_2, \quad a \cdot u_m \cdot u_3, \cdots, a \cdot u_m \cdot u_r \tag{2}$$

$$b \cdot u_m \cdot u_1, \quad b \cdot u_m \cdot u_2, \quad b \cdot u_m \cdot u_3, \cdots, b \cdot u_m \cdot u_r \tag{3}$$

where a and b belong to K. If any term of (2) is equal to any term of (3), then (2) and (3) are identical except for order. To show this suppose that $a \cdot u_m \cdot u_i = b \cdot u_m \cdot u_j$. Now u_j has an associate among the u's since they form a triplex. Hence

$$a \cdot u_m \cdot (u_i \cdot u'_j \cdot u_k) = b \cdot u_m \cdot u_j \cdot u'_j \cdot u_k = b \cdot u_m \cdot u_k.$$

Let k run from 1 to r. Then by Postulate *II* and Theorem 5 the expression $(u_i \cdot u'_j \cdot u_k)$ runs over all the u's without repetition. Hence (2) is some permutation of (3).

In (2) let a run over all the elements of K. From the resulting sequences let us save only those that are distinct and reject one of any pair of sequences which differ only in order. We are left then with a table of the form

$$\begin{array}{cccc} a_1 \cdot u_m \cdot u_1 & a_1 \cdot u_m \cdot u_2 & \cdots & a_1 \cdot u_m \cdot u_r \\ a_2 \cdot u_m \cdot u_1 & a_2 \cdot u_m \cdot u_2 & \cdots & a_2 \cdot u_m \cdot u_r \\ \cdot & \cdot & \cdot & \cdot \\ a_s \cdot u_m \cdot u_1 & a_s \cdot u_m \cdot u_2 & \cdots & a_s \cdot u_m \cdot u_r \end{array} \tag{4}$$

whose elements are distinct. In fact two equal elements could not be in different lines by the above reasoning, nor could they be in the same line since this would force two u's to be equal by Theorem 5. Let b be any element of K, then there is an element a for which

$$b = a \cdot u_m \cdot u_1.$$

Either a is one of the a_i $(i = 1, 2, \cdots, s)$ chosen above or else it was discarded in forming the table (4) because b had already occurred in the table. In either case b occurs in (4) and only once, since the elements of (4) are distinct. In short the elements of (4) are the elements of K. We have then

THEOREM 13. *The order of a sub-triplex is a proper divisor of the order of the triplex.*

As an immediate consequence of this theorem we have

COROLLARY 1. *The degree of any element of a triplex is a divisor of the order of the triplex.*

COROLLARY 2. *If a is any element of a triplex of order n, then $a^{2n+1} = a$.*

7. *The Quotient Triplex.* Let us return to the table (4) and discuss more of its properties.

THEOREM 14. *Let i, j, and k be any integers $\leqq s - n/r$. Then the set of r elements*

$$a_i \cdot a_j \cdot a_k \cdot u_m \cdot u_t \qquad (t = 1, 2, \cdots, r) \tag{5}$$

is, in some order, a line of the table (4).

Proof. Write $a = a_i \cdot a_j \cdot a_k$. Suppose $a \cdot u_m \cdot u_1$ lies in the ν-th line of (4). Then we can write

$$a \cdot u_m \cdot u_1 = a_\nu \cdot u_m \cdot u_\lambda \qquad (1 \leq \lambda \leq r).$$

If now u_ρ is any u, the corresponding member of (5) can be written

$$\begin{aligned} a \cdot u_m \cdot u_\rho &= a \cdot u_m \cdot (u_\rho \cdot u'_1 \cdot u_1) = (a \cdot u_m \cdot u_1) \cdot u_\rho \cdot u'_1 \\ &= a_\nu \cdot u_m \cdot (u_\lambda \cdot u_\rho \cdot u'_1) = a_\nu \cdot u_m \cdot u_\sigma \end{aligned}$$

hence $a \cdot u_m \cdot u_\rho$ is also in the ν-th line of the table and the theorem follows at once.

THEOREM 15. *If all the elements of any line of the table* (4) *be replaced by their associates, the result is a permutation of the elements of one of the lines of the table.*

Proof. Let us replace the elements of the ν-th line by their associates. Let the first element $a_\nu \cdot u_m \cdot u_1$ have its associate in the λ-th line so that $(a_\nu \cdot u_m \cdot u_1)' = a_\lambda \cdot u_m \cdot u_\omega$. Then the associate of any element in the ν-th line can be written

$$\begin{aligned}(a_\nu \cdot u_m \cdot u_\mu)' &= a'_\nu \cdot u'_m \cdot u'_\mu = a'_\nu \cdot u'_m \cdot u'_\mu \cdot u'_1 \cdot u_1 \\ &= (a_\nu \cdot u_m \cdot u_1)' \cdot u'_\mu \cdot u_1 = a_\lambda \cdot u_m \cdot (u_\omega \cdot u'_\mu \cdot u_1) \\ &= a_\lambda \cdot u_m \cdot u_\tau.\end{aligned}$$

Hence the associate lies also in the λ-th line. From this the theorem follows.

It becomes convenient to introduce a notation for the lines of the table (4). Accordingly we shall mean by l_i the i-th line or any permutation of its elements. We also define the result of operating by $\cdot$ on triplets of lines as follows

$$l_i \cdot l_j \cdot l_k = l_m,$$

where l_m is the line in which the element $a_i \cdot a_j \cdot a_k$ occurs. We further designate by l'_i the line containing the associates of the elements of l_i, and call either line the associate of the other. To show that this notation is proper we prove

THEOREM 16. *If l_k is any line of (4), then $l_i \cdot l'_i \cdot l_k = l_k$.*

Proof. Let $l'_i = l_j$. Then $l_i \cdot l'_i \cdot l_k \cdot l_i \cdot l_j$ is the line in which $a_i \cdot a_j \cdot a_k$ occurs. But a_j is the associate of some element of l_i. Hence we may write

$$a_j = (a_i \cdot u_m \cdot u_\rho)' = a'_i \cdot u'_m \cdot u'_\rho.$$

Therefore the line in question is that in which the element

$$a_i \cdot a'_i \cdot u'_m \cdot u'_\rho \cdot a_k = a_k \cdot u'_m \cdot u'_\rho \cdot u_m \cdot u'_m = a_k \cdot u_m \cdot (u'_m \cdot u'_m \cdot u'_\rho)$$

occurs. But $u'_m \cdot u'_m \cdot u'_\rho$ is some u, so that this element occurs in the k-th line and no other. Hence the theorem. We are now able to prove

THEOREM 17. *Under the operation $\cdot$ defined above the set of l's is a triplex.*

Proof. It is obvious that Postulate *I* is satisfied. As for Postulate *II* consider l_i, l_j and l_k as given. Then the expression $l_i' \cdot l_j' \cdot l_k = l_x$ is such that

$$l_i \cdot l_j \cdot l_x = l_i \cdot l'_i \cdot (l_j \cdot l'_j \cdot l_k) = l_j \cdot l'_j \cdot l_k = l_k.$$

Hence Postulate II is satisfied and the theorem follows.

It is clear that this triplex of lines is actually independent of the choice of a_i $(i = 1, 2, 3, \cdots, s)$. In fact some other choice would only permute

the elements of some of the lines. Furthermore the l's themselves do not actually depend on u_m for the same reason. Hence this triplex is uniquely determined by T and T_1 and corresponds to the quotient group G/H of G and its sub-group H.

THEOREM 18. *The quotient triplex is an extension of an Abelian group.*

Proof. The element u_1 must occur in the table (4). In fact there is a line l_μ which is made up of all the r elements of the sub-triplex T_1. Moreover a_μ is an element of T_1 as is also $a_\mu{}^3$. Hence $l_\mu \cdot l_\mu \cdot l_\mu = l_\mu$, so that $l'_\mu = l_\mu$, and by Theorem 8 the quotient triplex is an extension of an Abelian group.

8. *Cyclic Triplexes.* We proceed to give a few theorems concerning triplexes of a simple type. A triplex which consists of the distinct (odd) powers of a single element is called cyclic. In other words, a triplex is cyclic if and only if it contains an element whose degree is the order of the triplex. Such a triplex is simply isomorphic with the set of distinct odd powers of a primitive n-th root of unity, with the operation $\cdot$ as multiplication, or the set of odd integers $\leqq n$ under addition modulo n. From either of these points of view the following theorems are easily established.

THEOREM 19. *The order of any sub-triplex of a cyclic triplex contains exactly the same power of 2 as does the order of the triplex.*

THEOREM 20. *The sub-triplexes of a cyclic triplex are cyclic.*

THEOREM 21. *All cyclic triplexes of a fixed order n are simply isomorphic.*

THEOREM 22. *If n is the degree of an element a of a cyclic triplex of order n, then the degree of $b = a^\omega$ (ω odd) is $n/(n, \omega)$.*

THEOREM 23. *A cyclic triplex of order n contains exactly $\phi(2n)$ elements of degree n.*

Here ϕ is the totient function. More generally we have

THEOREM 24. *If $n = d\,\delta$ is any factorization of the order n of a cyclic triplex in which δ is odd, then there are exactly $\phi(2d)$ elements of the triplex of degree d.*

THEOREM 25. *No cyclic triplex of even order is an extension of a group.*

Proof. In fact the unit element of such a group would constitute a sub-triplex of order 1. This is impossible by Theorem 19.

THEOREM 26. *If the order of a cyclic triplex is a power of 2, then the triplex has no sub-triplexes.*

Proof. The existence of such a sub-triplex would contradict Theorem 19. As a converse of this theorem we have

THEOREM 27. *If a triplex T has no sub-triplex, then T is cyclic, and its order n is a power of 2.*

Proof. The degree of any element of T cannot be less than n, since otherwise we would have a sub-triplex of T. Hence T is cyclic. By Theorem 22 it follows that n must be prime to all odd numbers. That is, n is a power of 2.

9. *Examples.* We conclude with two simple examples of triplexes. The complete multiplication table of a triplex of order n is a cubical matrix of n^3 elements. On account of the symmetry of $a \cdot b \cdot c$ however, only one-sixth of this matrix need be given. But it is more convenient to form a plane table as follows

Example 1:

$$\begin{aligned} a &= a \cdot a \cdot b = a \cdot c \cdot d = b \cdot b \cdot b \cdot = b \cdot c \cdot c = b \cdot d \cdot d \\ b &= a \cdot a \cdot a \cdot = a \cdot b \cdot b = a \cdot c \cdot c = b \cdot c \cdot d = a \cdot d \cdot d \\ c &= a \cdot a \cdot d = a \cdot b \cdot c = b \cdot b \cdot d = c \cdot c \cdot d = d \cdot d \cdot d \\ d &= a \cdot a \cdot c = a \cdot b \cdot d = b \cdot b \cdot c = c \cdot c \cdot c = c \cdot d \cdot d. \end{aligned}$$

This triplex may also be described by

$$a^5 = a, \qquad c^5 = c, \qquad a \cdot a \cdot c = c \cdot c \cdot c.$$

Each letter is of degree 2 and $a' = b$, and $c' = d$. This triplex contains two cyclic sub-triplexes namely (a, b) and (c, d) both of which are simply isomorphic with the two letter triplex of § 3. The quotient triplex of each is the extension of two Abelian groups and may be described by

$$\begin{aligned} l_1 &= l_1 \cdot l_1 \cdot l_1 = l_1 \cdot l_2 \cdot l_2 \\ l_2 &= l_1 \cdot l_1 \cdot l_2 = l_2 \cdot l_2 \cdot l_2. \end{aligned}$$

Example 2:

$a^{13} = a$. This is a cyclic triplex of 6 letters. It contains only one sub-triplex namely (a^3, a^9). The quotient triplex is described by

$$l'_1 = l_1, \qquad l'_2 = l_3, \qquad l_2^5 = l_2.$$

STANFORD UNIVERSITY.

ON THE ASYMPTOTIC REPARTITION OF THE VALUES OF REAL ALMOST PERIODIC FUNCTIONS.

By Aurel Wintner.

In a previous paper * upon bounded matrices representing,† in the sense of Frobenius, the groups ‡ corresponding to the almost periodic functions of Bohr ¶ there is introduced the repartition function ("inverse function in the large") of a real almost periodic function. The presentation of the repartition function there given depends to some extent upon the theory of almost periodic matrices set forth in that paper. The aim of the present note is a more direct treatment which avoids any connection with the theory of matrices.—The methods are also valid for various generalizations ‖ of the notion of the almost-periodicity which are, however, for simplicity not treated in the present note.

Let f denote a real almost periodic function of the real argument t and m the largest, M the smallest number for which

$$m \leqq f(t) \leqq M, \qquad -\infty < t < +\infty. \tag{1}$$

Let

$$[f \leqq \xi]_T$$

denote the set of the points t for which

$$[f \leqq \xi]_T: \qquad -T \leqq t \leqq T \quad \text{and} \quad f(t) \leqq \xi \tag{2}$$

where ξ is any real number and T any positive number. The function **

---

* A. Wintner, "Diophantische Approximationen und Hermitesche Matrizen," *Mathematische Zeitschrift*, Vol. 30 (1929), pp. 290-319.

† These ideas have been inaugurated by O. Toeplitz, "Theorie der L-Formen," *Mathematische Annalen*, Vol. 70 (1911), p. 353; "Ueber das Wachstum der Potenzreihen in ihrem Konvergenzkreise. I," *Mathematische Zeitschrift*, Vol. 12 (1922), p. 195, and a not yet published paper on ordinary Dirichlet series.

‡ The abstract groups represented by the matrices which correspond to the limiting periodic (grenzperiodisch) subclass § of the class of the almost periodic functions have been studied recently by H. Schwerdtfeger, "Fastzyklische Gruppen," *Goettinger Nachrichten*, 1931, pp. 43-48.

§ H. Bohr, "Grenzperiodische Funktionen," *Acta Mathematica*, Vol. 52 (1928), pp. 127-133. Cf. G. D. Birkhoff, "On the Periodic Motions of Dynamical Systems," *Acta Mathematica*, Vol. 50 (1927), pp. 395-397.

¶ H. Bohr, "Zur Theorie der fast-periodischen Funktionen. I," *Acta Mathematica*, Vol. 45 (1925), pp. 29-127.

‖ Cf. H. Bohr, "Ueber die Verallgemeinerungen fastperiodischer Funktionen," *Mathematische Annalen*, Vol. 100 (1928), pp. 357-366.

** We denote by mes P the measure of the point set P.

$$(3) \qquad \rho_T(\xi) = (1/2T) \,.\, \mathrm{mes}\, [f \leqq \xi]_T$$

can be interpreted as yielding the relative frequency of the inequality $f(t) \leqq \xi$ in the range $-T \leqq t \leqq T$. In other words the difference

$$(4) \qquad \rho_T(\eta) - \rho_T(\xi) \quad \text{where} \quad \eta > \xi$$

represents the geometric probability that the number f lies in the strip $\xi < f \leqq \eta$ provided $-T \leqq t \leqq T$. The difference (4) is, according to (2) and (3), never negative, i. e. $\rho_T(\xi)$ is, for any fixed value of T, a monotone function of ξ for which we have, according to (1), (2), and (3),

$$(5) \quad \rho_T(\xi) \equiv 0 \text{ for } -\infty < \xi < m_T \text{ and } \rho_T(\xi) \equiv 1 \text{ for } M_T < \xi < +\infty$$

where

$$(6) \qquad \begin{aligned} m_T &= \min_{-T \leq t \leq T} f(t) \geqq \lim_{T=+\infty} m_T = m, \\ M_T &= \max_{-T \leq t \leq T} f(t) \leqq \lim_{T=+\infty} M_T = M. \end{aligned}$$

Since $f(t)$ is almost periodic and therefore continuous it follows from the definitions (1), (2), and (3) that the interval $m_T \leqq \xi \leqq M_T$ contains no subinterval in which $\rho_T(\xi)$ would be constant (this statement is simply the translation of the theorems of Bolzano and Weierstrass in the language of the repartition functions ρ). On the other hand the function $\rho_T(\xi)$ may have some discontinuity points (which may even lie everywhere dense in the interval $m_T \leqq \xi \leqq M_T$); this is, for instance, the case if the almost periodic function $f(t)$ is constant in some subintervals of the interval $-T \leqq t \leqq T$.— Since $f(t)$ is continuous it is also integrable in the sense of Lebesgue in any finite interval $-T \leqq t \leqq T$; the function (3) of ξ is obviously nothing more than the measure function (Massfunktion) of Lebesgue, belonging to the function $f(t)$ in the interval $-T \leqq t \leqq T$. It therefore follows, according to the theorems of Lebesgue * concerning the connection of his integral notion with that of Stieltjes, that

$$(7) \qquad (1/2T) \int_{-T}^{T} (f(t))^n \, dt = \int_{-\infty}^{+\infty} \xi^n \, d\rho_T(\xi); \qquad (n = 0, 1, 2, \cdots),$$

i. e.

$$(8) \qquad (1/2T) \int_{-T}^{T} (z - f(t))^{-1} \, dt = \int_{-\infty}^{+\infty} (z - \xi)^{-1} \, d\rho_T(\xi)$$

where z any constant (real or complex) which does not lie in the real segment ("Schlitz") $m_T \leqq z \leqq M_T$.

---

* H. Lebesgue, "Sur l'intégrale de Stieltjes et sur les opérations functionelles linéaires," *Comptes Rendus*, Vol. 150 (1910), pp. 86-88.

The results of the paper mentioned in the introduction which are independent of the theory of matrices concern an analogous *repartition theory belonging to the full definition domain* $-\infty < t < +\infty$ of the almost periodic function $f(t)$, i. e. *the (asymptotic) behaviour of the relative frequencies or probabilities in the large:* $T \rightarrow +\infty$. We shall write

$$(9) \qquad \mu(g) = \lim_{T=+\infty} (1/2T) \int_{-T}^{T} g(t)\, dt$$

provided this Hadamard average exists which is certainly the case if g is almost periodic and therefore if $g = f^n$ where $n = 0, 1, 2, \cdots$.—The results are as follows:

I. *There exists a monotone non-decreasing function* $\rho(\xi)$ *for which*

$$(10) \qquad \int_{-\infty}^{+\infty} \xi^n \, d\rho(\xi) = \mu(f^n); \qquad (n = 0, 1, 2, \cdots)$$

and the solution $\rho(\xi)$ *of* (10) *is uniquely determined if one normalizes it, for instance, by placing*

$$(11) \qquad \rho(-\infty) = 0; \qquad \rho(\xi) = \rho(\xi + 0), \qquad -\infty < \xi < +\infty;$$

i. e. the Hamburger momentum problem belonging to the sequences $\{\mu(f^n)\}$ is solvable and well-determined (déterminé).—This uniquely determined function $\rho(\xi)$ obviously possesses the usual statistical properties of a repartition function; from (10) and (9) there follows, for instance,

$$\mu(f^0) = \mu(1) = 1 = \int_{-\infty}^{+\infty} d\rho(\xi),$$

i. e. the total probability is equal to unity.

II. *The monotone function* $\rho(\xi)$ *which is uniquely determined by its formal properties* (10) *and* (11) *is effectively the asymptotic repartition function of the almost periodic function* $f(t)$. *In a more precise manner we have*

$$(12) \qquad \rho(\xi) = \lim_{T=+\infty} (1/2T) \,.\, \mathrm{mes}\, [f \leqq \xi]_T, \quad \text{i. e.} \quad \rho(\xi) = \lim_{T=+\infty} \rho_T(\xi)$$

for all continuity points of the function $\rho(\xi)$ which is monotone and has therefore at most a countable set of discontinuity points.—From (5), (6), and (12) there obviously follows

$$(13) \qquad \rho(\xi) \equiv 0 \text{ for } -\infty < \xi < m \text{ and } \rho(\xi) \equiv 1 \text{ for } M < \xi < +\infty.$$

III. *The function* $\rho(\xi)$ (which may have discontinuity points) *cannot*

be constant in an interval $\alpha < \xi < \beta$ *if* $m < \alpha$ *or* $\beta < M$ [cf. (1)]. This is the asymptotic refinement of the Bolzano-Weierstrass theorems mentioned above.

In the special case where the almost periodic function $f(t)$ is *periodic* and has the period τ it is clear that

$$(12') \qquad \rho_\tau(\xi) = \lim_{T=+\infty} \rho_T(\xi) \equiv \rho(\xi)$$

for *all* ξ, where T denotes as before a continuous parameter. In other words in the special case of periodic functions the limit (12) exists even if ξ is a discontinuity point of the asymptotic repartition function $\rho(\xi)$. In the paper mentioned in the introduction I was not able to show that the limit (12) also exists for the discontinuity points of the asymptotic repartition function $\rho(\xi)$ of any *almost* periodic function. Since that time it has been shown by Bohr * that not the apparatus applied by myself was at fault but that the limit (12) does not necessarily exist for an *almost* periodic function in the discontinuity points of its asymptotic repartition function. Since the theory of the Stieltjes transforms used by myself is accordingly the natural method † the present note applies, without modifications, for the demonstration of II the same apparatus.—One can start with the following theorem which follows readily from classical general results of Stieltjes and Helly ‡:

Let $\sigma_x(\xi)$ be a sheaf of functions of the variable ξ defined for $-\infty < \xi < +\infty$ and for all positive values of the parameter x of the sheaf. Suppose that $\sigma_x(\xi)$ is, for any fixed value of x, a monotone function of ξ:

$$(14a) \qquad \sigma_x(\eta) - \sigma_x(\xi) \geqq 0 \quad \text{for} \quad \eta > \xi$$

and that there exist two constants, c and C, independent of x, for which

$$(14b) \qquad \sigma_x(\xi) \equiv 0 \text{ for } -\infty < \xi < c \text{ and } \sigma_x(\xi) \equiv 1 \text{ for } C < \xi < +\infty;$$

suppose finally that there exists a domain § in the plane of the complex variable z in such a manner that the limit

$$(14c) \qquad \lim_{x=+\infty} \int_{-\infty}^{+\infty} (z-\xi)^{-1}\, d\sigma_x(\xi)$$

* H. Bohr, "Kleinere Beitraege zur Theorie der fastperiodischen Funktionen. II," *Det Kgl. Danske Videnskabernes Selskab. Meddelelser*, Vol. 10, Nr. 10, 1930.

† It is known that the Stieltjes-Helly theory is never able to yield any information at the discontinuity points of the limit function.

‡ Cf., for instance, A. Wintner, *Spektraltheorie der unendlichen Matrizen*, Leipzig, 1929, Chap. II.

§ By a domain we understand for simplicity a two-dimensional domain which does not contain any point of the real segment $c \leqq z \leqq C$ so that the integral behind the sign lim in (14c) obviously has a sense; cf. (14b).

exists for all points z of this domain. Then there exists in the range $-\infty < \xi < +\infty$ one and only one function $\sigma(\xi)$ for which

$$(14\alpha) \qquad \sigma(\xi) = \sigma(\xi + 0)$$

for all values of ξ and

$$(14\beta) \qquad \sigma(\xi) = \lim_{x=+\infty} \sigma_x(\xi)$$

for all continuity points of this function $\sigma(\xi)$ which is monotone (and has therefore at most a countable set of discontinuity points); furthermore, the function $\sigma(\xi)$, uniquely determined by (14α) and (14β), has the property that

$$(14\gamma) \qquad \lim_{x=+\infty} \int_{-\infty}^{+\infty} (z-\xi)^{-1} d\sigma_x(\xi) = \int_{-\infty}^{+\infty} (z-\xi)^{-1} d\sigma(\xi)$$

for all values of z which do not lie in the segment $c \leqq z \leqq C$; finally we have [cf. (14b), (14β)]

$$(14\delta) \quad \sigma(\xi) \equiv 0 \text{ for } -\infty < \xi < c \text{ and } \sigma(\xi) \equiv 1 \text{ for } C < \xi < +\infty$$

so that

$$(14\epsilon) \qquad \int_{-\infty}^{+\infty} = \int_{c-0}^{C+0}$$

in all occurring Stieltjes integrals.

Now if we put

$$(15) \qquad c = m, \qquad C = M, \qquad x = T, \qquad \sigma_x = \rho_T, \qquad \sigma = \rho$$

then condition (14a) is fulfilled inasmuch as the difference (4) is never negative. Condition (14b) is satisfied according to (5) and (6). Finally if z denotes any number which does not lie in the segment $m \leqq z \leqq M$ of the z-plane then $g(t) = (z - f(t))^{-1}$ is, according to (1), an almost periodic function so that expression (8) approaches for $T = +\infty$ the Hadamard average (9):

$$(15c) \qquad \lim_{T=+\infty} \int_{-\infty}^{+\infty} (z-\xi)^{-1}\, d\rho_T(\xi) = \mu((z-f)^{-1}).$$

All three conditions (14a), (14b), (14c) are accordingly satisfied.—The proof of I, II, and III (p. 341) proceeds now as follows.

Proof of I. On denoting by $\rho(\xi)$ the function $\sigma(\xi)$ which is uniquely defined by (14α) and (14β) equations (14γ), (15c), and (15) yield

$$(16) \qquad \int_{-\infty}^{+\infty} (z-\xi)^{-1}\, d\rho(\xi) = \mu((z-f)^{-1})$$

for all values of z which do not lie in the segment $m \leqq z \leqq M$ and therefore certainly for all values of z which do not lie in the circle $|z| \leqq R$ where

$$(17) \qquad R \geqq |m|, \qquad R \geqq |M|.$$

Since the series

$$(z - f(t))^{-1} = \sum_{n=0}^{\infty} z^{-n-1}(f(t))^n$$

is, according to (1) and (17), uniformly convergent with respect to t for any fixed value of z for which $|z| > R$, we have

$$(18) \qquad \mu((z-f)^{-1}) = \sum_{n=0}^{\infty} z^{-n-1}\mu(f^n) \quad \text{for} \quad |z| > R.$$

From (14ϵ), (15), and (17) there follows in an analogous manner

$$(19) \qquad \int_{-\infty}^{+\infty} (z-\xi)^{-1}\, d\rho(\xi) = \sum_{n=0}^{\infty} z^{-n-1} \int_{-\infty}^{+\infty} \xi^n\, d\rho(\xi) \quad \text{for} \quad |z| > R.$$

On introducing (18) and (19) in (16) and comparing the coefficients one obtains the momentum equations (10) the solution $\rho(\xi) = \sigma(\xi)$ of which fulfills, according to (14α) and (14δ), the conditions (11). In order to finish the proof of I we show that the solution $\rho(\xi)$ found before is unique. The solution $\rho(\xi)$ of (10) and (11) is, according to Carleman,* certainly unique if

$$(20) \qquad \sum_{n=0}^{\infty} \left[\int_{-\infty}^{+\infty} \xi^{2n}\, d\rho(\xi) \right]^{-1/2n} = +\infty.$$

Equation (20) is an obvious consequence of (14ϵ). In fact, we have for any number r which is larger than $|c|$ and $|C|$

$$0 < \int_{-\infty}^{+\infty} \xi^{2n}\, d\rho(\xi) = \int_{-r}^{r} \xi^{2n}\, d\rho(\xi) < r^{2n} \int_{-r}^{r} d\rho(\xi) = r^{2n}$$

which clearly is sufficient for (20).

Proof of II. Equation (12) and (13) follow, by virtue of (15), from (14β) and (14δ), respectively.

Proof of III. The function $f(t)$ is, by supposition, almost periodic,

---

* T. Carleman, "Sur les séries asymptotiques," *Comptes Rendus*, Vol. 174 (1922), pp. 1527-1530, and "Sur les functions quasi-analytiques," *Wissenschaftliche Vortraege, gehalten am 5. Kongress der Skandinavischen Mathematiker*, Helsingfors, 1923, pp. 186-187.

i. e. it is continuous in every point of the range $-\infty < t < +\infty$ and has the property that there exists, for any fixed $\epsilon > 0$, a sequence

$$\cdots, \; l_{-2}(\epsilon), \; l_{-1}(\epsilon), \; l_0(\epsilon), \; l_1(\epsilon), \; l_2(\epsilon), \cdots$$

of numbers $l_\nu = l_\nu(\epsilon)$ and a number $K = K(\epsilon)$, independent of ν, for which

$$(21) \qquad 0 < l_{\nu+1}(\epsilon) - l_\nu(\epsilon) < K(\epsilon), \qquad (\nu = 0, \pm 1, \pm 2, \cdots)$$

$$\lim_{\nu=+\infty} l_\nu(\epsilon) = +\infty, \qquad \lim_{\nu=-\infty} l_\nu(\epsilon) = -\infty$$

and

$$(22) \qquad | f(t + l_\nu(\epsilon)) - f(t) | < \epsilon \quad \text{for} \quad -\infty < t < +\infty, \quad (\nu = 0, \pm 1, \pm 2, \cdots).$$

Let $\xi = \alpha$ be any ***inner*** point of the range $m \leqq \xi \leqq M$ of $f(t)$. Then there exists a number t_0 for which

$$(23) \qquad f(t_0) = \alpha.$$

After α and t_0 are fixed one can determine, by virtue of the continuity of $f(t)$, for any sufficiently small $\delta > 0$ a number $\tau = \tau(\delta) > 0$ for which

$$(24) \qquad | f(t) - \alpha | < \delta \quad \text{if} \quad t_0 - \tau(\delta) < t < t_0 + \tau(\delta),$$

inasmuch as α is an inner point $(m < \alpha < M)$. Conditions (22), (23), and (24) yield

$$(25) \quad | f(t + l_\nu(\delta)) - f(t_0) | < 2\delta \quad \text{for} \quad t_0 - \tau(\delta) < t < t_0 + \tau(\delta), \quad (\nu = 0, \pm 1, \pm 2, \cdots).$$

According to (25) and (21) it is clearly * impossible that the limit (12) should be a constant function of ξ in the vicinity of the point $\xi = \alpha$.

It is not possible to obtain further results for the asymptotic repartition function of the most general real-valued almost-periodic function. However, in the important special case where the frequencies of the almost-periodic functions are linearly-independent, the method of momentums applied above leads to more refined and explicit results. Cf. a paper to appear in the *Mathematische Zeitschrift.*

THE JOHNS HOPKINS UNIVERSITY.

---

* Analogous considerations have often been used by Bohr; cf., for instance, ***Acta Mathematica***, Vol. 45 (1925), p. 55.

ON A GENERALIZATION OF THE LAGRANGE-GAUSS MODULAR ALGORITHM.

By Aurel Wintner.

Introduction. In the present note the iteration process

$$(1) \qquad a_{n+1} = 2^{-1}(a_n + b_n), \quad b_{n+1} = (a_n b_n)^{1/2}; \qquad (n = 0, 1, 2, \cdots),$$
$$a_0 = a > 0, \quad b_0 = b > 0$$

of the arithmetical-geometrical mean which is the Landen transformation rule of the elliptic modular functions and has furnished the base of the Gaussian theory of elliptic functions will be generalized as follows: a_{n+1} should be not $2^{-1}(a_n + b_n)$ but

$$a_{n+1} = (2^{-1}(a_n^{\rho} + b_n^{\rho}))^{1/\rho}$$

where ρ is any positive number which may be even irrational.

First it will be shown that both limits

$$\lim_{n=\infty} a_n, \qquad \lim_{n=\infty} b_n$$

exist and have a common value which will be denoted by $\mathrm{M}_\rho(a, b)$ inasmuch as it is a function of a, b, and ρ. The value of $\mathrm{M}_1(a, b)$ which belongs to (1) is known to be

$$(1') \qquad \mathrm{M}_1(a, b) = 2^{-1}\pi\left[\int_0^{\pi/2}(a^2\cos^2\phi + b^2\sin^2\phi)^{-1/2}\,d\phi\right]^{-1}.$$

The more general limit $\mathrm{M}_\rho(a, b)$ for which I have not found an analogous explicit representation is also a "mean value" contained in the interval (a, b). The limit $\mathrm{M}_\rho(a, b)$ is then studied as a function of the exponent ρ. Finally the asymptotic behavior of $\mathrm{M}_\rho(a, b)$ for $\lim(b/a) = 0$ is established. There is obtained a direct generalization of the fundamental result of Gauss according to which

$$(1'') \qquad \lim_{b/a=0} \mathrm{M}_1(a, b)\log(4a/b) = a\pi/2$$

and which may be derived from the hypergeometric differential equation of Legendre:

$$\alpha = 1/2, \qquad \beta = 1/2, \qquad \gamma = 1$$

defining the periods of the complete elliptic integral (1′). It is, however, not certain that there exists a hypergeometric differential equation (with more general indices α, β, γ) for which the generalized process ($\rho > 0$) could yield

an analogous explicit Grenzkreisuniformisierung.—The considerations necessary in the case $\rho > 0$ are in some points more complicated than those furnishing the classical theory $(\rho = 1)$.

The existence of the generalized Gaussian mean. If A and σ are positive numbers then we understand by A^σ the real branch of the power A^σ so that $A^\sigma > 0$.—First we show that if

$$(2) \quad a_{n+1} = (2^{-1}(a_n{}^\rho + b_n{}^\rho))^{1/\rho}, \quad b_{n+1} = (a_n b_n)^{1/2}; \qquad (n = 0, 1, 2, \cdots)$$

where $\rho > 0$ and

$$(2') \qquad a_0 = a > 0, \qquad b_0 = b > 0$$

then a_n and b_n approach a common limit

$$(3) \qquad \mathrm{M}_\rho(a, b) = \lim_{n=\infty} a_n = \lim_{n=\infty} b_n.$$

Since a_n and b_n are for $a = b$ independent of n we may suppose that

$$(4') \qquad a \neq b.$$

We show that $\mathrm{M}_\rho(a, b)$ lies, for all values of a, b, and ρ, between a and b:

$$(4) \qquad \min(a, b) < \mathrm{M}_\rho(a, b) < \max(a, b).$$

Since the algorithm (2) is homogeneous we have

$$(5) \qquad \lambda \mathrm{M}_\rho(a, b) = \mathrm{M}_\rho(\lambda a, \lambda b); \qquad (\lambda > 0).$$

The function $\mathrm{M}_\rho(a, b)$ fulfills, according to (2) and (3), the functional equation

$$(6) \qquad \mathrm{M}_\rho(a, b) = \mathrm{M}_\rho([2^{-1}(a^\rho + b^\rho)]^{1/\rho}, (ab)^{1/2})$$

which is, of course, a characteristic property of the limit (3).

First we deduce from (2) the identities

$$(7) \qquad 2(a^\rho{}_{n+1} + b^\rho{}_{n+1}) = (a_n{}^{\rho/2} + b_n{}^{\rho/2})^2$$

and

$$(8) \qquad 4(a_{n+1}{}^{2\rho} - b_{n+1}{}^{2\rho}) = (a_n{}^\rho - b_n{}^\rho)^2$$

the second of which may be written in the form

$$(9) \quad 4(a^\rho{}_{n+1} - b^\rho{}_{n+1}) : (a_n{}^\rho - b_n{}^\rho) = (a_n{}^{\rho/2} + b_n{}^{\rho/2})(a_n{}^{\rho/2} - b_n{}^{\rho/2}) : (a^\rho{}_{n+1} + b^\rho{}_{n+1}).$$

On comparing (9) with (7) one obtains

$$2(a^\rho{}_{n+1} - b^\rho{}_{n+1}) : (a_n{}^\rho - b_n{}^\rho) = (a_n{}^{\rho/2} - b_n{}^{\rho/2}) : (a_n{}^{\rho/2} + b_n{}^{\rho/2}),$$

so that

$$| 2(a^\rho_{n+1} - b^\rho_{n+1}) : (a_n^\rho - b_n^\rho) | < 1,$$

i. e.

$$| a^\rho_{n+1} - b^\rho_{n+1} | < 2^{-1} | a_n^\rho - b_n^\rho | ;$$

hence, according to (2′),

$$(10) \qquad | a_n^\rho - b_n^\rho | < 2^{-n} | a^\rho - b^\rho | .$$

Furthermore there follows from (8), (4′), (2) and (2′) the inequality

$$a_{n+1}{}^{2\rho} - b_{n+1}{}^{2\rho} > 0,$$

i. e., by virtue of $\rho > 0$, simply

$$(11a) \qquad a_{n+1} > b_{n+1}; \qquad (n = 0, 1, 2, \cdots)$$

so that, according to (2),

$$(11b) \qquad a_{n+2} = (2^{-1}(a^\rho_{n+1} + b^\rho_{n+1}))^{1/\rho} < (2^{-1}(a^\rho_{n+1} + a^\rho_{n+1}))^{1/\rho} = a_{n+1}$$

and

$$(11c) \qquad b_{n+2} = (a_{n+1} b_{n+1})^{1/2} > (b_{n+1} b_{n+1})^{1/2} = b_{n+1}.$$

On collecting (11a), (11b) and (11c) we obtain the inequalities

$$(11) \qquad b_n < b_{n+1} < a_{n+1} < a_n; \qquad (n = 1, 2, \cdots)$$

according to which both sequences

$$a_1, a_2, \cdots; \quad b_1, b_2, \cdots$$

are monotone and bounded so that the limits $\lim a_n$, $\lim b_n$ necessarily exist. The more precise statement (3) is an obvious consequence of (10). Since, according to (2) and (4′),

$$(11d) \qquad \begin{aligned} a_1 &= (2^{-1}(a^\rho + b^\rho))^{1/\rho} < \max(a, b), \\ b_1 &= (ab)^{1/2} > \min(a, b) \end{aligned}$$

we have, according to (11),

$$(11') \qquad \min(a, b) < a_n < b_n < \max(a, b).$$

From (11), (11′) and (3) there results the inequality (4).

Continuity of the function $\mathbf{M}_\rho$. We now show that $\mathbf{M}_\rho(a, b)$ is a continuous function of the pair (a, b). It follows readily from (5) that it is sufficient to prove the continuity of $\mathbf{M}_\rho(1, b)$ which is a function of the single variable b and may be represented, according to (3), by the everywhere $(b > 0)$ convergent series

$$(12) \qquad b_1 + \sum_{n=1}^{\infty} (b_{n+1} - b_n).$$

The terms of this series are, according to (2), (2′), (11) and (11′), positive continuous functions of b alone ($a = 1$). From our inequalities there obviously follows the uniform convergence of (12) in any finite range

$$0 < b < \text{const.} \tag{12'}$$

On using complex arguments a, b it would be easy to prove that $M_\rho(a, b)$ is not only continuous but also analytic.

The mean $M_\rho(a, b)$ *as a function of* ρ. If a and b have fixed values then the numbers (2) are functions of ρ only, say

$$a_n = a_n(\rho), \qquad b_n = b_n(\rho)$$

In order to show that

$$a_n(\rho) < a_n(\sigma) \quad \text{for} \quad \rho < \sigma;\ n > 1 \tag{13}$$

and

$$b_n(\rho) < b_n(\sigma) \quad \text{for} \quad \rho < \sigma;\ n > 1 \tag{14}$$

we notice the well-known inequality

$$(2^{-1}(A^\rho + B^\rho))^{1/\rho} < (2^{-1}(A^\sigma + B^\sigma))^{1/\sigma} \quad \text{for} \quad \rho < \sigma \tag{15}$$

which is valid for any pair of positive numbers A, B, provided $A \neq B$. For

$$A = a, \qquad B = b$$

we obtain from (15)

$$a_1(\rho) < a_1(\sigma) \tag{13'}$$

and from (11d) there follows

$$b_1(\rho) = b_1(\sigma). \tag{14'}$$

Since $a_1(\rho)$ is, according to (11), different from $b_1(\rho)$ we may apply inequality (15) for

$$A = a_1(\rho), \qquad B = b_1(\rho)$$

and obtain

$$a_2(\rho) = (2^{-1}(a_1(\rho)^\rho + b_1(\rho)^\rho)^{1/\rho} < (2^{-1}(a_1(\rho)^\sigma + b_1(\rho)^\sigma)^{1/\sigma},$$

hence, according to (13′) and (14′),

$$a_2(\rho) < (2^{-1}(a_1(\sigma)^\sigma + b_1(\sigma)^\sigma))^{1/\sigma},$$

i. e.

$$a_2(\rho) < a_2(\sigma). \tag{13a}$$

Since, according to (2),

$$b_2(\rho) = (a_1(\rho)b_1(\rho))^{1/2}, \qquad b_2(\sigma) = (a_1(\sigma)b_1(\sigma))^{1/2},$$

we have, from (13′) and (14′),

(14a) $$b_2(\rho) < b_2(\sigma).$$

Inequalities (13) and (14) accordingly hold for $n = 2$. On proceeding in this manner we obtain that (13) and (14) are generally valid.

There follows now from (13), (14) and (3) that $\mathbf{M}_\rho(a, b)$ *is a monotone function of* ρ. Since $\mathbf{M}_\rho(a, b)$ is, according to (4), a bounded function of ρ we conclude that *the limits*

(16) $$\mathbf{M}_0(a, b) = \lim_{\rho=0} \mathbf{M}_\rho(a, b), \qquad \mathbf{M}_\infty(a, b) = \lim_{\rho=\infty} \mathbf{M}_\rho(a, b)$$

exist.

The retrograde Legendre chain of the complementary moduli for the means $\mathbf{M}_\rho$. We now suppose [cf. (4′)] that

(17) $$a > b \ (> 0)$$

and put

(18) $$c = (a^{2\rho} - b^{2\rho})^{1/2\rho} = c_0$$

and in a more general manner

(19) $$c_n = (a_n{}^{2\rho} - b_n{}^{2\rho})^{1/2\rho}; \qquad (n = 0, 1, 2, \cdots).$$

We have then from (2′), (17), (11) and (11′)

(20) $$0 < b_n < b_{n+1} < a_{n+1} < a_n; \qquad (n = 0, 1, 2, \cdots)$$

and from (19)

(21) $$c_n > 0; \qquad (n = 0, 1, 2, \cdots).$$

From (19) and (8) there follows

(22) $$4c_{n+1}{}^{2\rho} = 4(a_{n+1}{}^{2\rho} - b_{n+1}{}^{2\rho}) = (a_n{}^\rho - b_n{}^\rho)^2$$

i. e.

(23) $$c_{n+1} = (2^{-1}(a_n{}^\rho - b_n{}^\rho))^{1/\rho}.$$

On substituting the values (2) in the following formulae [which result from (2) and (23)]:

(23′) $$a_{n+2} = (2^{-1}(a^\rho{}_{n+1} + b^\rho{}_{n+1}))^{1/\rho}, \quad c_{n+2} = (2^{-1}(a^\rho{}_{n+1} - b^\rho{}_{n+1}))^{1/\rho}$$

one obtains after an easy reduction the identities

(24) $$a_{n+2}{}^{\rho/2} = 2^{-1}(a_n{}^{\rho/2} + b_n{}^{\rho/2}), \quad c_{n+2}{}^{\rho/2} = 2^{-1}(a_n{}^{\rho/2} - b_n{}^{\rho/2})$$

according to which

(25) $$a_{n+2}{}^{\rho/2} + c_{n+2}{}^{\rho/2} = a_n{}^{\rho/2}.$$

Furthermore, we have from (19)

$$c_n^{2\rho} = a_n^{2\rho} - b_n^{2\rho} = (a_n^\rho + b_n^\rho)(a_n^\rho - b_n^\rho),$$

so that, according to (2) and (23),

(26) $$c_n^{2\rho} = 4a^\rho_{n+1}\, c^\rho_{n+1}.$$

There follows finally from (2) and (23) the relation

(27) $$a^\rho_{n+1} + c^\rho_{n+1} = a_n^\rho.$$

Since, according to (27) and (26),

$$a_{n-1} = 2^{1/\rho}\,[2^{-1}(a_n^\rho + c_n^\rho)]^{1/\rho}, \quad c_{n-1} = 2^{1/\rho}\,[(a_n\, c_n)^{1/2}]$$

the functional equations (5) and (6) yield

$$\mathrm{M}_\rho(a_{n-1}, c_{n-1}) = 2^{1/\rho}\,\mathrm{M}_\rho(a_n, c_n)$$

so that, according to (2′) and (18),

(28) $$\mathrm{M}_\rho(a, c) = 2^{n/\rho}\,\mathrm{M}_\rho(a_n, c_n); \qquad (n = 0, 1, 2, \cdots).$$

On the other hand, from (2) and (6) there follows simply

(29) $$\mathrm{M}_\rho(a, b) = \mathrm{M}_\rho(a_n, b_n); \qquad (n = 0, 1, 2, \cdots).$$

Since all numbers a_n, b_n, c_n are positive, we may put

(30) $$h_n = 2^{-n\rho} \log\,(2^{2/\rho}\, a_n/c_n); \qquad (n = 0, 1, 2, \cdots)$$

so that [cf. (18)]

(30′) $$h_0 = \log\,(2^{2/\rho}/[1 - (b/a)^{2\rho}]^{1/2\rho})$$

where we take the real branch of the logarithm. Since, from (26),

(26′) $$1/c_{n+1} = 2^{2/\rho}\, a_{n+1}/c_n^2$$

equation (30) yields

$$h_{n+1} = 2^{-(n+1)\rho} \log\,(2^{2/\rho}\, a_{n+1}/c_{n+1}) = 2^{-(n+1)\rho} \log\,[(2^{2/\rho}\, a_{n+1}/c_n)^2]$$

i. e.

(31′) $$2^{\rho-1}\, h_{n+1} = 2^{-n\rho} \log\,(2^{2/\rho} a_{n+1}/c_n),$$

so that, by virtue of (30),

(31) $$2^{\rho-1}\, h_{n+1} - h_n = 2^{-n\rho} \log\,(a_{n+1}/a_n); \qquad (n = 0, 1, 2, \cdots).$$

On solving the recursion formula (31) by complete induction one obtains the explicit representation

(32) $$2^{(\rho-1)n}\, h_n = h_0 + \sum_{m=1}^{n-1} 2^{-m} \log\,(a_{m+1}/a_m)$$

which assures the existence of the limit

(33) $$\Lambda = \lim_{n=\infty} 2^{(\rho-1)n} h_n = h_0 + \sum_{m=1}^{\infty} 2^{-m} \log (a_{m+1}/a_m)$$

inasmuch as

$$0 < \lim_{m=\infty} a_m = \lim_{m=\infty} a_{m+1}, \quad \text{so that} \quad \lim_{m=\infty} \log (a_{m+1}/a_m) = 0.$$

On solving the recursion formulae (24) and (25) by complete induction and taking the limit for $n = \infty$ Gauss has obtained, in his case $\rho = 1$, the power development of the elliptic theta functions.

The asymptotic behavior of $\mathbf{M}_\rho(1, \kappa)$ *for* $\lim \kappa = 0$. We may now write equation (28), by means of the homogeneity property (5), in the form

(34) $$2^{-n} = [\mathbf{M}_\rho(1, c_n/a_n) : \mathbf{M}_\rho(a/a_n, c/a_n]^\rho$$

and substitute the identities (30) and (34) in the limiting relation (33) obtaining

(35) $$\Lambda = \lim_{n=\infty} \{[\mathbf{M}_\rho(1, c_n/a_n) : \mathbf{M}_\rho(a/a_n, c/a_n)]^\rho \log (2^{3/\rho} a_n/c_n)\}.$$

Since we have, by virtue of (3) and the continuity of the function $\mathbf{M}_\rho$,

$$\lim_{n=\infty} \mathbf{M}_\rho(a/a_n, c/a_n) = \mathbf{M}_\rho(a/\mathbf{M}_\rho(a, b), c/\mathbf{M}_\rho(a, b))$$

i. e. [cf. (5)]

(36) $$\lim_{n=\infty} \mathbf{M}_\rho(a/a_n, c/a_n) = \mathbf{M}_\rho(a, c) : \mathbf{M}_\rho(a, b)$$

so that the relation (35) is equivalent with

(37) $$[\mathbf{M}_\rho(a, b) : \mathbf{M}_\rho(a, c)]^{-\rho} \Lambda = \lim_{n=\infty} \{[\mathbf{M}_\rho(1, c_n/a_n)]^\rho \log (2^{3/\rho} a_n/c_n)\}$$

where, according to (3), (4), and (19),

$$\lim_{n=\infty} c_n = 0, \qquad \lim_{n=\infty} a_n = \mathbf{M}(a, b) > 0.$$

Since the ratio c_n/a_n may take, according to (3), (2) and (19), *any* sufficiently small positive value we may write

$$p_\rho = \lim_{\epsilon=0} \{\mathbf{M}_\rho(1, \epsilon) [\log (2^{3/\rho} \epsilon^{-1})]^\rho\}$$

where ϵ is a *continuous* independent variable. *This relation determines the order of vanishing of the function* $\mathbf{M}_\rho(1, \kappa)$ *of the independent variable* κ *for* $\lim \kappa = +0$. The numerical value of the number p_ρ is known only for $\rho = 1$, viz.

$$p_1 = \pi/2$$

[cf. (1″)].

THE JOHNS HOPKINS UNIVERSITY.

ON THE DISTRIBUTIONS OF CERTAIN STATISTICS.*

By Allen T. Craig.

Introduction. It is the purpose of this paper to present certain general theories on the distributions of such statistics as the arithmetic mean, harmonic mean, geometric mean, median, quartile, decile and range of samples of n items selected at random from a rather arbitrary universe. We shall also develop, without appeal to n-dimensional geometry the distribution of the Pearson χ^2. Exact distributions are known only for a few statistics and these in case the sampled universes are of special types. For a normal parent population, the distribution of the means of samples of n was probably known to Gauss, while Czuber † seems first to have obtained the distribution of the sum of the squares of the deviations of the n items of the sample from the sample mean. Czuber's derivation was closely related to that of Helmert ‡ who dealt with the distribution of sums of squares of deviations from a fixed point. These results of Helmert and Czuber were apparently unknown to Student § who found the distributions of the standard deviation and of the ratio of the deviation of the mean of the sample from the population mean to the standard deviation of the sample in case the parent population is normal. Student's results, somewhat empirically obtained, were later verified by Karl Pearson.¶ By geometrical methods, R. A. Fisher also verified Student's distribution and obtained the distributions of the correlation coefficient,‖ the correlation ratio,** the regression coefficient,** the partial correlation coefficient †† and the multiple correlation coefficient ‡‡

* Presented to the Society, December 31, 1930.

† E. Czuber, *Theorie der Beobachtungsfehler* (1891), pp. 161-162.

‡ Helmert, "Über die Wahrscheinlichkeit der Potenzsummen," *Zeitschrift für Mathematik und Physik*, Bd. 21 (1876).

§ Student, "The Probable Error of a Mean," *Biometrika*, Vol. 6 (1908), pp. 1-25.

¶ K. Pearson, "Standard Deviations of Small Samples," *Biometrika*, Vol. 10 (1914-15), pp. 522-529.

‖ R. A. Fisher, "Frequency Distribution of the Values of the Correlation Coefficient, etc.," *Biometrika*, Vol. 10 (1914-15), pp. 507-521.

** R. A. Fisher, "The Goodness of Fit of Regression Formulae and the Distribution of Regression Coefficients," *Journal of the Royal Statistical Society*, Vol. 85 (1922).

†† R. A. Fisher, "The Distribution of the Partial Correlation Coefficient," *Metron*, Vol. 3 (1924), pp. 329-333.

‡‡ R. A. Fisher, "The General Sampling Distribution of the Multiple Correlation Coefficient," *Proceedings of the Royal Society*, Vol. 121 (1928).

for normal universes. Much of Fisher's work has been given analytic treatment by Romanovsky.* With the exception of the arithmetic mean, very little is known regarding the exact distributions of statistics obtained from samples taken from non-normal universes. Church † has shown that arithmetic means of samples of n items drawn from a universe represented by a Pearson Type III curve are distributed in accordance with a curve of the same type. This result was later found by Irwin ‡ and C. C. Craig.§ In the same paper, Irwin gave the distribution of arithmetic means of samples of n from a rectangular universe, a result obtained simultaneously by Hall.¶ Both Irwin's and Hall's work on the rectangular universe was considerably antedated by Laplace ‖ and Rietz,** who gave the distribution of the sums of n items from this type of universe. Irwin †† later extended his method of integral equations to obtain the distributions of means of samples of n items drawn from universes represented by Types I and VII of the Pearson system of frequency curves. Baker ‡‡ has treated the distributions of the arithmetic mean and standard deviation for certain non-homogeneous populations and in a later paper §§ extended the method to a universe which can be represented by a finite number of terms of a Gram-Charlier series.

1. *The probability function.* In the present paper, we shall understand a probability function $f(x)$ of a real variable x to be, for all values of x on a range X, a single-valued, non-negative, continuous function with

* V. Romanovsky, "On the Moments of Standard Deviations, etc.," ***Metron***, Vol. 5 (1925), Part 4, pp. 3-45; "On the Criteria That Two Given Samples Belong to the Same Normal Population," ***Metron***, Vol. 7 (1928), Part 3, pp. 3-46; "On the Distribution of the Regression Coefficient, etc.," ***Bulletin de l'Académie des Sciences de l'U. R. S. S.*** (1926).

† A. E. R. Church, "On the Means and Squared Standard Deviations, etc.," ***Biometrika***, Vol. 18 (1926), pp. 321-394.

‡ J. O. Irwin, "On the Frequency Distributions of Means, etc.," ***Biometrika***, Vol. 19 (1927), pp. 225-239.

§ C. C. Craig, "Sampling When the Parent Population is of Pearson's Type III," ***Biometrika***, Vol. 21 (1929), pp. 287-289.

¶ P. Hall, "The Distribution of Means for Samples of Size N, etc.," ***Biometrika***, Vol. 19 (1927), pp. 240-245.

‖ Laplace, ***Théorie Analytique des Probabilités***, Troisieme Ed., (1820), pp. 257-263.

** H. L. Rietz, "On a Certain Law of Probability of Laplace," ***Proceedings of the International Mathematical Congress at Toronto*** (1924), pp. 795-799.

†† J. O. Irwin, "On the Frequency Distributions of Means, etc.," ***Metron***, Vol. 8 (1930), pp. 51-105.

‡‡ G. A. Baker, "Random Sampling from Non-Homogeneous Populations," ***Metron***, Vol. 8 (1930), pp. 67-87.

§§ G. A. Baker, "Distribution of the Means, etc.," ***Annals of Mathematical Statistics***, Vol. 1 (1930), pp. 199-204.

$\int_X f(x)\,dx = 1$. Then $\int_a^b f(x)\,dx$ is the probability that a value of x selected at random lies in interval (a, b) where a and b are in X and $a < b$; and $f(x)\,dx$ is, to within infinitesimals of order higher than that of dx, the probability that a value of x selected at random lies in the interval $(x, x + dx)$. It will prove convenient to classify probability functions according as X is the range $(-\infty, \infty)$, $(0, \infty)$, or $(0, A)$, $A > 0$. In accord with this classification, and by adopting the language of Bachelier * to some extent, we shall refer to probability functions as of the first, second, and third kinds respectively.

2. *The distribution of the arithmetic mean.*

THEOREM I.† *Let $f(x)$ be the probability function of a variable x. Let $F(y)$ be that of the sum $y = \Sigma t_i$ where t_i $(i = 1, 2, \cdots n)$ are n independent values of x. If $f(x)$ is a probability function of the first kind, $F(y)$ is a probability function of the first kind, and*

$$(1) \qquad F(y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(t_1)f(t_2) \cdots f(t_{n-1})f(y - t_1 - \cdots - t_{n-1})\,dt_{n-1}\cdots dt_1.$$

Moreover, the probability function $\phi_1(\bar{x})$ of $\bar{x} = y/n$ is a probability function of the first kind and

$$(1') \qquad \phi_1(\bar{x}) = nF(n\bar{x}).$$

Proof. By the definition of $F(y)$, $\int_y^{y+\Delta y} F(t)\,dt$ is the probability that y lies in the interval $(y, y + \Delta y)$. For y assigned, it is clear that $t_1, t_2, \cdots, t_{n-1}$ may be chosen arbitrarily from X whereas t_n must be selected with certainty from the interval $(y - t_1 - \cdots - t_{n-1},\ y + \Delta y - t_1 - \cdots - t_{n-1})$ in order that the sum may lie in the interval $(y, y + \Delta y)$. The probability of the joint occurrence of these events is then the product of the probabilities of the separate occurrences integrated over the prescribed intervals. Hence,

$$\int_y^{y+\Delta y} F(t)\,dt = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\int_{y-t_1-\cdots-t_{n-1}}^{y+\Delta y-t_1-\cdots-t_{n-1}} f(t_1)f(t_2)\cdots f(t_n)\,dt_n\cdots dt_1.$$

Integrate both members of the equation and denote the primitive of a function by writing a bar over it. We have

---

* L. Bachelier, *Calcul des Probabilités*, p. 155.

† Cf. R. von Mises, "Fundamentalsätze der Wahrscheinlichkeitsregnung," *Mathematische Zeitschrift*, Vol. 4 (1919), pp. 20, 21, 77. Also C. V. L. Charlier, "Contributions to the Mathematical Theory of Statistics," *Arkiv for Matematik, Astronomi öch Fysik*, Bd. 8, No. 4 (1912), pp. 7-9.

$$\bar{F}(y+\Delta y)-\bar{F}(y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}f(t_1)\cdots f(t_{n-1})$$
$$\times[\bar{f}(y+\Delta y-t_1-\cdots-t_{n-1})-\bar{f}(y-t_1-\cdots-t_{n-1})]\,dt_{n-1}\cdots dt_1.$$

Next, divide both members by Δy and pass to the limit as Δy approaches zero. This establishes the main part of the theorem. Since $F(y)dy$ is, to within infinitesimals of higher order, the probability that a value of y chosen at random lies in the interval $(y, y+dy)$, the change of variable $y=n\bar{x}$ yields the probability function $nF(n\bar{x})$ of $\bar{x}$.

If $f(x)$ is a probability function of the second kind, it is clear that for y assigned, we may choose t_1 arbitrarily from the interval $(0, y)$; t_2 from the interval $(0, y-t_1)$; $\cdots$ and t_{n-1} from the interval $(0, y-t_1-\cdots-t_{n-2})$. Finally, t_n must be selected with certainty from the interval

$$(y-t_1-\cdots-t_{n-1},\ y+\Delta y-t_1-\cdots t_{n-1})$$

in order that the sum may lie in the interval $(y, y+\Delta y)$. Accordingly, for a probability function of the second kind, (1) takes the form

$$(1.1)\qquad F(y)=\int_0^y\int_0^{y-t_1}\cdots\int_0^{y-t_1-t_2\cdots-t_{n-2}}f(t_1)\cdots f(t_{n-1})$$
$$\times f(y-t_1-\cdots-t_{n-1})\,dt_{n-1}\cdots dt_1$$

while (1′) is unaltered.

We find it convenient to treat the case of a probability function of the third kind in more detail. Let us suppose $f(x)$ to be defined for values of x on the range $(0, a)$. We shall first exhibit $F(y)$ for sums of $n=2, 3, 4$ values. The limits of integration may be easily verified. In this manner we avoid a lengthy exposition for sums of n values. We have for $n=2$,*

$$(1.2)\qquad F(y)=\int_0^y f(t_1)f(y-t_1)\,dt_1,\qquad (0\leq y\leq a),$$
$$=\int_{y-a}^a f(t_1)f(y-t_1)\,dt_1,\qquad (a\leq y\leq 2a);$$

for $n=3$,†

$$(1.3)\qquad F(y)=\int_0^y\int_0^{y-t_1}f(t_1)f(t_2)f(y-t_1-t_2)\,dt_2dt_1,\qquad (0\leq y\leq a),$$
$$=\left[\int_0^{y-a}\int_{y-a-t_1}^a+\int_{y-a}^a\int_0^{y-t_1}\right]f(t_1)f(t_2)f(y-t_1-t_2)\,dt_2dt_1,$$
$$(a\leq y\leq 2a),$$

* Cf. Czuber, *loc. cit.*, p. 69. Also K. Mayr, "Wahrscheinlichkeitsfunktionen und ihre Anwendungen," *Monatshefte für Mathematik und Physik*, Band 30 (1920), p. 20.

† Czuber, *loc. cit.*, p. 73.

$$= \int_{y-2a}^{a} \int_{y-a-t_1}^{a} f(t_1) f(t_2) f(y - t_1 - t_2)\, dt_2 dt_1, \qquad (2a \leq y \leq 3a);$$

for $n = 4$,

$$(1.4) \qquad F(y) = \int_0^y \int_0^{y-t_1} \int_0^{y-t_1-t_2} \Theta(y, t_1, t_2, t_3)\, dt_3 dt_2 dt_1, \qquad (0 \leq y \leq a)$$

where $\Theta(y, t_1, t_2, t_3) = f(t_1) f(t_2) f(t_3) f(y - t_1 - t_2 - t_3)$

$$= \Big[\int_0^{y-a} \int_0^{y-a-t_1} \int_{y-a-t_1-t_2}^{a} + \int_0^{y-a} \int_{y-a-t_1}^{a} \int_0^{y-t_1-t_2} + \int_{y-a}^{a} \int_0^{y-t_1} \int_0^{y-t_1-t_2} \Big] \Theta(y, t_1, t_2, t_3)\, dt_3 dt_2 dt_1, \qquad (a \leq y \leq 2a),$$

$$= \Big[\int_0^{y-2a} \int_{y-2a-t_1}^{a} \int_{y-a-t_1-t_2}^{a} + \int_{y-2a}^{a} \int_0^{y-a-t_1} \int_{y-a-t_1-t_2}^{a} + \int_{y-2a}^{a} \int_{y-a-t_1}^{a} \int_0^{y-t_1-t_2} \Big] \Theta(y, t_1, t_2, t_3)\, dt_3 dt_2 dt_1, \qquad (2a \leq y \leq 3a).$$

The results of extending the forms of expressions of $F(y)$ in terms of integrals, which we have obtained for $n = 2, 3, 4$, to higher values of n now becomes fairly obvious. For example, with sums of k values, $F(y)$ is defined on $(0, a)$ by a $(k-1)$-fold integral; on $(a, 2a)$ by the sum of $\binom{k-1}{1}$ $(k-1)$-fold integrals; and in general on $[ra, (r+1)a]$ by the sum of $\binom{k-1}{r}$ such integrals. With this in mind and by employing the following abbreviated notations

$$\Theta(y, t_1, \cdots, t_{n-1}) = f(t_1) f(t_2) \cdots f(y - t_1 - \cdots - t_{n-1}),$$
$$x_{ij} = y - ia - t_1 - \cdots - t_j,$$

$$\int_0^y \int_0^{y-t_1} \cdots \int_0^{y-t_1-t_2-\cdots t_{n-2}} \Theta(y, t_1, \cdots, t_{n-1})\, dt_{n-1} \cdots dt_1$$
$$= \begin{pmatrix} x_{00} & x_{01} & x_{02} & \cdots & x_{0,n-2} \\ 0 & 0 & 0 & & 0 \end{pmatrix} \Theta(y, t_1, \cdots, t_{n-1})\, dt_{n-1} \cdots dt_1,$$

we readily see that for sums of n values

$$(1.5) \quad F(y) = \begin{pmatrix} x_{00} & x_{01} & x_{02} & \cdots & x_{0,n-2} \\ 0 & 0 & 0 & & 0 \end{pmatrix} \Theta(y, t_1, \cdots t_{n-1})\, dt_{n-1} \cdots dt_1, \qquad (0 \leq y \leq a),$$

$$
\begin{aligned}
&-\Bigg[\begin{pmatrix} x_{10} & x_{11} & \cdots & x_{1,n-3} & a \\ 0 & 0 & & 0 & x_{1,n-2} \end{pmatrix} \\
&\quad + \begin{pmatrix} x_{10} & x_{11} & \cdots & x_{1,n-4} & a & x_{0,n-2} \\ 0 & 0 & & 0 & x_{1,n-3} & 0 \end{pmatrix} \\
&\quad + \begin{pmatrix} x_{10} & x_{11} & \cdots & x_{1,n-5} & a & x_{0,n-3} & x_{0,n-2} \\ 0 & 0 & & 0 & x_{1,n-4} & 0 & 0 \end{pmatrix} \\
&\quad + \\
&\quad \vdots \\
&\quad + \begin{pmatrix} x_{10} & x_{11} & a & x_{03} & \cdots & x_{0,n-2} \\ 0 & 0 & x_{12} & 0 & & 0 \end{pmatrix} \\
&\quad + \begin{pmatrix} x_{10} & a & x_{02} & \cdots & x_{0,n-2} \\ 0 & x_{11} & 0 & & 0 \end{pmatrix} \\
&\quad + \begin{pmatrix} a & x_{01} & \cdots & x_{0,n-2} \\ x_{10} & 0 & & 0 \end{pmatrix}\Bigg] \Theta(y, t_1, \cdots, t_{n-1})\, dt_{n-1} \cdots dt_1, \\
&\hspace{22em} (a \leqq y \leqq 2a), \\
&-\Bigg[\begin{pmatrix} x_{20} & x_{21} & \cdots & x_{2,n-4} & a & a \\ 0 & 0 & & 0 & x_{2,n-3} & x_{1,n-2} \end{pmatrix} \\
&\quad + \begin{pmatrix} x_{20} & x_{21} & \cdots & x_{2,n-5} & a & x_{1,n-3} & a \\ 0 & 0 & & 0 & x_{2,n-4} & 0 & x_{1,n-2} \end{pmatrix} \\
&\quad + \begin{pmatrix} x_{20} & x_{21} & \cdots & x_{2,n-6} & a & x_{1,n-4} & x_{1,n-3} & a \\ 0 & 0 & & 0 & x_{2,n-5} & 0 & 0 & x_{1,n-2} \end{pmatrix} \\
&\quad + \\
&\quad \vdots \\
&\quad + \begin{pmatrix} x_{20} & x_{21} & a & x_{13} & \cdots & x_{1,n-3} & a \\ 0 & 0 & x_{22} & 0 & & 0 & x_{1,n-2} \end{pmatrix} \\
&\quad + \begin{pmatrix} x_{20} & a & x_{12} & \cdots & x_{1,n-3} & a \\ 0 & x_{21} & 0 & & 0 & x_{1,n-2} \end{pmatrix} \\
&\quad + \begin{pmatrix} a & x_{11} & x_{12} & \cdots & x_{1,n-3} & a \\ x_{20} & 0 & 0 & & 0 & x_{1,n-2} \end{pmatrix} \\
&\quad + \begin{pmatrix} x_{20} & x_{21} & \cdots & x_{2,n-5} & a & a & x_{0,n-2} \\ 0 & 0 & & 0 & x_{2,n-4} & x_{1,n-3} & 0 \end{pmatrix} \\
&\quad + \begin{pmatrix} x_{20} & x_{21} & \cdots & x_{2,n-6} & a & x_{1,n-4} & a & x_{0,n-2} \\ 0 & 0 & & 0 & x_{2,n-5} & 0 & x_{1,n-3} & 0 \end{pmatrix} \\
&\quad + \\
&\quad \vdots
\end{aligned}
$$

$$+\begin{pmatrix} x_{20} & x_{21} & a & x_{13} & \cdots & x_{1,n-4} & a & x_{0,n-2} \\ 0 & 0 & x_{22} & 0 & & 0 & x_{1,n-3} & 0 \end{pmatrix}$$

$$+\begin{pmatrix} x_{20} & a & x_{12} & \cdots & x_{1,n-4} & a & x_{0,n-2} \\ 0 & x_{21} & 0 & & 0 & x_{1,n-3} & 0 \end{pmatrix}$$

$$+\begin{pmatrix} a & x_{11} & \cdots & x_{1,n-4} & a & x_{0,n-2} \\ x_{20} & 0 & & 0 & x_{1,n-3} & 0 \end{pmatrix}$$

$$+$$

$$\vdots$$

$$+\begin{pmatrix} x_{20} & a & a & x_{03} & \cdots & x_{0,n-2} \\ 0 & x_{21} & x_{12} & 0 & & 0 \end{pmatrix}$$

$$+\begin{pmatrix} a & x_{11} & a & x_{03} & \cdots & x_{0,n-2} \\ x_{20} & 0 & x_{12} & 0 & & 0 \end{pmatrix}$$

$$+\begin{pmatrix} a & a & x_{02} & \cdots & x_{0,n-2} \\ x_{20} & x_{11} & 0 & & 0 \end{pmatrix}\Bigg] \Theta(y, t_1, \cdots, t_{n-1})\, dt_{n-1} \cdots dt_1, \qquad (2a \leq y \leq 3a),$$

$$\vdots$$

$$= \begin{pmatrix} a & a & \cdots & a \\ x_{n-1,0} & x_{n-2,1} & & x_{1,n-2} \end{pmatrix} \Theta(y, t_1, \cdots, t_{n-1})\, dt_{n-1} \cdots dt_1, \qquad [(n-1)\, a \leq y \leq na].$$

As in the preceding cases, the distribution function of the arithmetic mean $\bar{x} = y/n$ is obtained from that of y by means of (1′).

As applications of the foregoing theorems, we shall now find in as simple algebraic form as possible the distribution functions of the sums and arithmetic means of n items drawn from various types of simple non-normal parent populations.

1. Consider $$f(x) = 1/a, \qquad (0 \leq x \leq a).$$

This is a very special case of the theory, and as reference has been made in the introduction to the distribution function of sums of n items drawn from this type of universe, we shall not take the space to give the results.

2. Consider $$f(x) = (k/\sigma) e^{-x/\sigma}, \qquad (0 \leq x < \infty).$$

By (1.1), $$F(y) = [k^n y^{n-1}/\sigma^n (n-1)\,!]\; e^{-y/\sigma}$$

and $$\phi_1(\bar{x}) = [(kn)^n \bar{x}^{n-1}/\sigma^n (n-1)\,!] e^{-n\bar{x}/\sigma}, \qquad (0 \leq \bar{x} < \infty).$$

3. Consider $$f(x) = kx^{-1/2} e^{-x/2}, \qquad (0 \leq x < \infty),$$

which is the distribution of the squares of items distributed normally * about their expected value in the population. The distribution of sums of n items drawn from this population will then be of special interest as it is the distribution of the Pearson χ^2. Again, by (1.1) we have

$$F(y) = \int_0^y \int_0^{y-t_1} \cdots \int_0^{y-t_1-\cdots-t_{n-2}} (k/t_1^{1/2})(k/t_2^{1/2})$$

$$\cdots k/(y-t_1-t_2-\cdots-t_{n-1})^{1/2} e^{-y/2}\, dt_{n-1}\cdots dt_1$$

$$= \pi k^n e^{-y/2} \int_0^y (dt_1/t_1^{1/2}) \int_0^{y-t_1} (dt_2/t_2^{1/2}) \cdots \int_0^{y-t_1-t_2-\cdots-t_{n-3}} (dt_{n-2}/t_{n-2}^{1/2})$$

$$= 2\pi k^n e^{-y/2} \int_0^y (dt_1/t_1^{1/2}) \int_0^{y-t_1} (dt_2/t_2^{1/2})$$

$$\cdots \int_0^{y-t_1-\cdots-t_{n-4}} (y-t_1-\cdots-t_{n-3})(dt_{n-3}/t_{n-3}^{1/2}).$$

When integrating with respect to t_{n-s}, $s=3,4,\cdots,n-1$, make the transformation

$$z = [(y-t_1-\cdots-t_{n-s-1})/t_{n-s}]^{1/(s-2)}.$$

The integral considered becomes

$$\int_1^\infty (y-t_1-\cdots-t_{n-s})^{(s-1)/2}\psi(z)\,dz$$

where $\int_1^\infty \psi(z)\,dz$ exists. Accordingly, we find

$$F(y) = k_1 y^{(n-2)/2} e^{-y/2}$$

and

$$\phi_1(\bar{x}) = k_2 \bar{x}^{(n-2)/2} e^{-n\bar{x}/2}, \qquad (0 \leq \bar{x} < \infty).$$

4. Consider

$$f(x) = (k/\sigma) e^{-|x|/\sigma} \qquad (-\infty < x < \infty).$$

For $n=2$,

$$\phi_1(\bar{x}) = (2k^2/\sigma^2)(\sigma + 2\,|\,\bar{x}\,|\,) e^{-2|\bar{x}|/\sigma}, \qquad (-\infty < \bar{x} < \infty).$$

For $n=3$,

$$\phi_1(\bar{x}) = (9k^3/2\sigma^3)(\sigma^2 + 3\sigma\,|\,\bar{x}\,|\, + 3\bar{x}^2) e^{-3|\bar{x}|/\sigma}, \qquad (-\infty < \bar{x} < \infty).$$

For $n=4$,

$$\phi_1(\bar{x}) = (2k^4/3\sigma^4)(15\sigma^3 + 60\sigma^2\,|\,\bar{x}\,|\, + 98\sigma\bar{x}^2 + 64\,|\,\bar{x}^3\,|\,) e^{-4|\bar{x}|/\sigma}, \qquad (-\infty < \bar{x} < \infty).$$

* H. L. Rietz, "Frequency Distributions Obtained by Certain Transformations of Normally Distributed Variables," *Annals of Mathematics*, Second Series, Vol. 23 (1921-22), p. 294.

5. Consider

$$f(x) = 4x/a^2, \qquad (0 \leq x \leq a/2),$$
$$= (4/a^2)(a - x), \qquad (a/2 \leq x \leq a).$$

For $n = 2$,

$$\phi_1(\bar{x}) = (2^8/3!\,a^4)\bar{x}^3, \qquad (0 \leq \bar{x} \leq a/4),$$
$$= (2^4/3!\,a^4)(-48\bar{x}^3 + 48a\bar{x}^2 - 12a^2\bar{x} + a^3), \quad (a/4 \leq \bar{x} \leq a/2),$$
$$= (2^4/3!\,a^4)(48\bar{x}^3 - 96a\bar{x}^2 + 60a^2\bar{x} - 11a^3), \quad (a/2 \leq \bar{x} \leq 3a/4),$$
$$= (2^8/3!\,a^4)(-\bar{x}^3 + 3a\bar{x}^2 - 3a^2\bar{x} + a^3), \qquad (3a/4 \leq \bar{x} \leq a).$$

It is interesting to note that $f(x)$ represents the distribution of means of samples of two and that $\phi_1(\bar{x})$ represents the distribution of means of samples of four for a rectangular parent distribution.

6. Consider

$$f(x) = 2x/a^2, \qquad (0 \leq x \leq a).$$

For $n = 2$,

$$\phi_1(\bar{x}) = (2^5/3a^4)\bar{x}^3, \qquad (0 \leq \bar{x} \leq a/2),$$
$$= (2^4/3a^4)(-2\bar{x}^3 + 3a^2\bar{x} - a^3), \qquad (a/2 \leqq \bar{x} \leqq a).$$

For $n = 3$,

$$\phi_1(\bar{x}) = (3^6/5a^6)\bar{x}^5, \qquad (0 \leq \bar{x} \leq a/3),$$
$$= (3/5a^6)(-162\bar{x}^5 + 270a^2\bar{x}^3 - 180a^3\bar{x}^2 + 45a^4\bar{x} - 4a^5), \qquad (a/3 \leq \bar{x} \leq 2a/3),$$
$$= (3^3/5a^6)(9\bar{x}^5 - 30a^2\bar{x}^3 + 20a^3\bar{x}^2 + 5a^4\bar{x} - 4a^5), \qquad (2a/3 \leq \bar{x} \leq a).$$

2. *The distribution of the harmonic mean.* Let $f(x)$ be the probability function of x. It is well known * that the probability function of $x' = 1/x$ is $F(x') = (1/x'^2)f(1/x')$ provided $1/x$ is continuous on the interval on which $f(x)$ is defined. By use of the theorems on probability functions for sums, we may find $h(y')$, the probability function of $y' = \Sigma\, t'_i$ where t'_i are independent values of x'. The change of variable $y = y'/n$ yields $\psi(y)$, the probability function of the reciprocal of the harmonic mean of samples drawn from the universe represented by $f(x)$. The harmonic mean $\bar{\bar{x}} = 1/y$ then has for its probability function

$$\phi_2(\bar{\bar{x}}) = (1/\bar{\bar{x}}^2)\psi(1/\bar{\bar{x}}). \tag{2}$$

---

* E. L. Dodd, "The Frequency Law of a Function of One Variable," *Bulletin of the American Mathematical Society*, Vol. 31 (1925), p. 28. Also, "The Frequency Law of a Function of Variables With Given Frequency Laws," *Annals of Mathematics*, Second Series, Vol. 27 (1925-26), p. 18; K. Mayr, *loc. cit.*, p. 20.

As an example, let us consider

$$f(x) = 2a/x^2, \qquad (a \leq x \leq 2a).$$

We find

$$F(x') = 2a, \qquad (1/2a \leq x' \leq 1/a).$$

For $n = 2$,

$$\begin{aligned} h(y') &= 4a^2(y' - 1/a), && (1/a \leq y' \leq 3/2a), \\ &= 4a^2(2/a - y'), && (3/2a \leq y' \leq 2/a), \\ \psi(y) &= 16a^2(y - 1/2a), && (1/2a \leq y \leq 3/4a), \\ &= 16a^2(1/a - y), && (3/4a \leq y \leq 1/a), \end{aligned}$$

and

$$\begin{aligned} \phi_2(\bar{\bar{x}}) &= (16a^2/\bar{\bar{x}}^2)(1/a - 1/\bar{\bar{x}}), && (a \leq \bar{\bar{x}} \leq 4a/3), \\ &= (16a^2/\bar{\bar{x}}^2)(1/\bar{\bar{x}} - 1/2a), && (4a/3 \leq \bar{\bar{x}} \leq 2a). \end{aligned}$$

For $n = 3$,

$$\begin{aligned} \phi_2(\bar{\bar{x}}) &= (3^2 4!\, a^3/\bar{\bar{x}}^2)(1/2\bar{\bar{x}}^2 - 1/a\bar{\bar{x}} + 1/2a^2), && (a \leq \bar{\bar{x}} \leq 6a/5), \\ &= (4!\, a^3/\bar{\bar{x}}^2)(-9/\bar{\bar{x}}^2 + 27/2a\bar{\bar{x}} - 39/8a^2), && (6a/5 \leq \bar{\bar{x}} \leq 3a/2), \\ &= (3^2 4!\, a^3/\bar{\bar{x}}^2)(1/2\bar{\bar{x}}^2 - 1/2a\bar{\bar{x}} + 1/8a^2), && (3a/2 \leq \bar{\bar{x}} \leq 2a). \end{aligned}$$

Also consider

$$f(x) = 1/a, \qquad (a \leq x \leq 2a).$$

For $n = 2$,

$$\phi_2(\bar{\bar{x}}) = [2/a^2]\,[(\bar{\bar{x}}/2)\log(\bar{\bar{x}}/[2a - \bar{\bar{x}}]) + a\bar{\bar{x}}/2(2a - \bar{\bar{x}}) - a/2], \qquad (a \leqq \bar{\bar{x}} \leqq 4a/3),$$

$$= [2/a^2]\,[(\bar{\bar{x}}/2)\log(4a - \bar{\bar{x}}/\bar{\bar{x}}) - a\bar{\bar{x}}/(4a - \bar{\bar{x}}) + a], \qquad (4a/3 \leq \bar{\bar{x}} \leq 2a).$$

3. *The distribution of the geometric mean.*

THEOREM II. *Let $f(x)$ be the probability function of the variable x and let $F(y)$ be the probability function of the variable product $y = \Pi\, t_i$ where t_i $(i = 1, 2, \cdots, n)$ are n independent values of x. If $f(x)$ is a probability function of the second kind, $F(y)$ is a probability function of the second kind and*

$$(3) \quad F(y) = \int_0^\infty \int_0^\infty \cdots \int_0^\infty \frac{f(t_1)f(t_2)\cdots f(t_{n-1})f[y/(t_1\cdots t_{n-1})]}{t_1\cdots t_{n-1}}\, dt_{n-1}\cdots dt_1.$$

Moreover, the probability function $\phi_3(z)$ of $z = (y)^{1/n}$ is a probability function of the second kind and

$$(3') \qquad \phi_3(z) = nz^{n-1}F(z^n).$$

Proof. For y assigned, we may select $t_1, t_2, \cdots, t_{n-1}$ arbitrarily from

the interval $(0, \infty)$. But t_n must be selected with certainty from the interval $[y/(t_1 \cdots t_{n-1}),\ (y + \Delta y)/(t_1 \cdots t_{n-1})]$ in order that the product may lie in $(y, y + \Delta y)$. We then have

$$\int_y^{y+\Delta y} F(t)\,dt = \int_0^\infty \cdots \int_0^\infty \int_{y/(t_1 \cdots t_{n-1})}^{(y+\Delta y)/(t_1 \cdots t_{n-1})} f(t_1) \cdots f(t_n)\,dt_n \cdots dt_1$$

or

$$\bar{F}(y + \Delta y) - \bar{F}(y) = \int_0^\infty \cdots \int_0^\infty f(t_1) \cdots f(t_{n-1})$$
$$\times \{\bar{f}\,[(y + \Delta y)/(t_1 \cdots t_{n-1})] - \bar{f}\,[y/(t_1 \cdots t_{n-1})]\}\,dt_{n-1} \cdots dt_1.$$

If we assume the validity of the expansion

$$\bar{f}\left(\frac{y}{t_1 \cdots t_{n-1}} + \frac{\Delta y}{t_1 \cdots t_{n-1}}\right) = \bar{f}\left(\frac{y}{t_1 \cdots t_{n-1}}\right) + f'\left(\frac{y}{t_1 \cdots t_{n-1}}\right)\frac{\Delta y}{t_1 \cdots t_{n-1}} + \cdots,$$

the theorem is proved.

If $f(x)$ is a probability function of the third kind, we establish in a precisely similar manner that the probability function $F(y)$ of the product of n independent values of the variable x is

$$F(y) = \int_{y/a^{n-1}}^{a} \int_{y/a^{n-2}t_1}^{a} \cdots \int_{y/(at_1 \cdots t_{n-2})}^{a} \frac{f(t_1) \cdots f(t_{n-1}) f\,[y/(t_1 \cdots t_{n-1})]}{t_1 \cdots t_{n-1}}\,dt_{n-1} \cdots dt_1. \tag{3.1}$$

Clearly, the distribution of geometric means is given by (3′).

Inasmuch as the geometric mean of n numbers is well-defined only when all of the numbers are positive, this completes our theory of the distribution of this average. As an example, consider

$$f(x) = 1/a, \qquad (0 \leq x \leq a).$$

For $n = 2$,

$$\phi_2(z) = (4z/a^2) \log (a/z), \qquad (0 \leq z \leq a).$$

For $n = 3$,

$$\phi_3(z) = (27z^2/2a^3)\,[\log (a/z)]^2, \qquad (0 \leq z \leq a).$$

4. *The distributions of the median,* quartile and decile.* In a paper by Rider,† the distribution of the medians of samples of $2m + 1$ items is given in the case of a rectangular parent distribution. We shall extend the theory to a more general distribution.

Suppose a variable x to obey a law of probability given by $f(x)$. For the

* Cf. Dodd, "Functions of Measurements under General Laws of Error," *Skandinavisk Aktuarietidskrift* (1922), p. 150.

† P. R. Rider, "On the Distribution of the Ratio of Mean to Standard Deviation in Small Samples from Non-Normal Universes," *Biometrika*, Vol. 21 (1929), pp. 136-137.

present we shall assume $f(x)$ to be of the third kind. Let a sample consisting of $n = 2m + 1$ (m an integer) values of x with median value ξ be drawn. We propose to determine the law of probability of ξ. The probability that m of the $2m + 1$ items lie in the interval $(0, \xi)$ is $\binom{2m+1}{m}\left[\int_0^\xi f(t)\,dt\right]^m$. The probability that m of the remaining $m + 1$ items lie in the interval (ξ, a) is $\binom{m+1}{m}\left[\int_\xi^a f(t)\,dt\right]^m$. Finally, the probability that the remaining item lies in $(\xi, \xi + d\xi)$ is $f(\xi)\,d\xi$. We have thus established the following

THEOREM III. *Let $f(x)$ be the probability function of the variable x and let $\phi_4(\xi)$ be the probability function of the median ξ of samples of $n = 2m + 1$. If $f(x)$ is a probability function of the third kind, $\phi_4(\xi)$ is a probability function of the third kind and*

$$\phi_4(\xi) = [(2m+1)!/(m!)^2]\left[\int_0^\xi f(t)\,dt\right]^m\left[\int_\xi^a f(t)\,dt\right]^m f(\xi). \tag{4}$$

If $f(x)$ is a probability function of the first or second kind, $\phi_4(\xi)$ remains unchanged except that the limits of integration are $-\infty, \xi$ and ξ, ∞ respectively for a function of the first kind and $0, \xi$ and ξ, ∞ respectively for a function of the second kind.

The probability function of the lower quartile η of samples of $n = 4m + 1$ drawn from a universe represented by $f(x)$ where $f(x)$ is of the second kind, is clearly

$$\phi_5(\eta) = [(4m+1)!/m!(3m)!]\left[\int_0^\eta f(t)\,dt\right]^m\left[\int_\eta^\infty f(t)\,dt\right]^{3m} f(\eta). \tag{4.1}$$

Similar forms of $\phi_5(\eta)$ hold for probability functions of the first and third kinds. It is obvious that the distribution of any statistic which is defined as that value of the variate which exceeds, and is exceeded by, specified numbers of items of the sample, may be determined in this manner. For example, we may find the distribution of the upper quartile or of any decile.

Examples. Consider

$$f(x) = (k/\sigma)e^{-|x|/\sigma}, \qquad (-\infty < x < \infty).$$

Then

$$\phi_4(\xi) = [(2m+1)!/(m!)^2]\,(k^{2m+1}/\sigma)e^{-(1/\sigma)(m+1)|\xi|}(2 - e^{-(1/\sigma)|\xi|})^m, \qquad (-\infty < \xi < \infty).$$

Consider $\qquad f(x) = (2/a^2)x, \qquad (0 \leq x \leq a).$

Then

$$\phi_4(\xi) = [(2m+1)!/(m!)^2]\,(2/a^{4m+2})\xi^{2m+1}(a^2 - \xi^2)^m, \qquad (0 \leq \xi \leq a).$$

Consider $\qquad f(x) = e^{-x}, \qquad (0 \leq x < \infty).$

Then

$$\phi_4(\xi) = [(2m+1)!/(m!)^2]\, e^{-(m+1)\xi}(1-e^{-\xi})^m, \qquad (0 \leq \xi < \infty).$$

Consider * $\qquad f(x) = 1/a, \qquad (0 \leq x \leq a).$

Then

$$\phi_4(\xi) = [(2m+1)!/(m!)^2] \cdot (1/a^{2m+1})\xi^m(a-\xi)^m, \qquad (0 \leq \xi \leq a).$$

It is interesting to compare the variances of these exact distributions of medians with the variances given by the usual formula † $\sigma_m{}^2 = 1/4Ny^2$ where N is the number in the sample and y is the ordinate of the probability (not frequency) function of the population at the population median. For the population represented by $f(x) = 1/a$, $\sigma_m{}^2 = a^2/4(2m+3)$. By the usual formula, $\sigma_m{}^2 = a^2/4(2m+1)$. For the population represented by $f(x) = 2x/a^2$, we have

$$\sigma_m{}^2 = \frac{a^2}{2} - \frac{a^2\,[(2m+1)!]^6\, 2^{4m+4}}{(m!)^4\,[(4m+3)!]^2},$$

while by the ordinary formula $\sigma_m{}^2 = a^2/8(2m+1)$. It is easy to verify that the limit of

$$\frac{[(2m+1)!]^6\, 2^{4m+4}}{(m!)^4\,[(4m+3)!]^2}$$

is 1/2 as m becomes infinite.

5. *The distribution of the range.* In their investigation of test criteria, Neyman and Pearson ‡ found, for a rectangular universe, that the ranges of samples of n items are distributed according to a Pearson Type I curve. In this connection, we shall establish the slightly more general

THEOREM IV. *Let $f(x)$ be the probability function of the variable x and let $\phi_6(W)$ be the probability function of the range W of samples of n items. If $f(x)$ is a probability function of the third kind, $\phi_6(W)$ is a probability function of the third kind and*

$$(5) \qquad \phi_6(W) = n(n-1)\int_W^a f(x)f(x-W)\left[\int_{x-W}^{x} f(x)\,dx\right]^{n-2} dx.$$

Proof. Consider a W assigned within an interval $(W, W+\Delta W)$. The largest item of the sample must be at least as large as W. If x is this largest item, then in order that the range W may lie in the interval $(W, W+\Delta W)$, it is necessary that the smallest item lie in $(x-W-\Delta W,\ x-W)$. The remaining $n-2$ items must lie in the interval $(x-W, x)$. If we consider

* Rider, *loc. cit.*, p. 137.

† H. L. Rietz, *Mathematical Statistics* (1927), p. 135.

‡ J. Neyman and E. S. Pearson, "On the Use and Interpretation of Certain Test Criteria," *Biometrika*, Vol. 20 (1928), p. 210.

all possible arrangements and denote by $\int_W^{W+\Delta W} \phi_6(W)_x dW$ the probability, for x assigned to the interval $(x, x+dx)$, that W lies in the interval $(W, W+\Delta W)$, we have

$$\int_W^{W+\Delta W} \phi_6(W)_x dW = n(n-1) \left[\int_{x-W}^{x} f(x)\,dx\right]^{n-2} f(x)\,dx \int_{x-W-\Delta W}^{x-W} f(x)\,dx$$

whence

$$\phi_6(W)_x = n(n-1) \left[\int_{x-W}^{x} f(x)\,dx\right]^{n-2} f(x) f(x-W)\,dx.$$

Now x may take all values from W to a. Hence, the probability function of W is

$$\phi_6(W) = n(n-1) \int_W^a f(x) f(x-W) \left[\int_{x-W}^{x} f(x)\,dx\right]^{n-2} dx.$$

For the probability functions of the first and second kinds, we find

$$\phi_6(W) = n(n-1) \int_{-\infty}^{\infty} f(x) f(x-W) \left[\int_{x-W}^{x} f(x)\,dx\right]^{n-2} dx,$$

and

$$\phi_6(W) = n(n-1) \int_W^{\infty} f(x) f(x-W) \left[\int_{x-W}^{x} f(x)\,dx\right]^{n-2} dx$$

respectively.

If a sample consists of two items, the standard deviation is $s = \frac{1}{2} \mid t_1 - t_2 \mid$ or $s = \frac{1}{2} W$. We may, accordingly, obtain the probability function $\phi_7(s)$ of the standard deviation of samples of two drawn from an arbitrary universe by setting $n = 2$ in the above forms of $\phi_6(W)$ and by making the transformation $W = 2s$, $dW = 2ds$.

Examples. Consider *

$$f(x) = 1/a, \qquad (0 \leq x \leq a).$$

We find

$$\phi_6(W) = [n(n-1)/a^{n-1}] W^{n-2}(1 - W/a), \qquad (0 \leq W \leq a),$$

and

$$\phi_7(s) = (4/a^2)(a - 2s), \qquad (0 \leq s \leq a/2).$$

Consider

$$f(x) = e^{-x}, \qquad (0 \leq x < \infty).$$

We find

$$\phi_6(W) = (n-1) e^{(1-n)W}(e^W - 1)^{n-2}, \qquad (0 \leq W < \infty),$$

and

$$\phi_7(s) = 2e^{-2s}, \qquad (0 \leq s < \infty).$$

THE UNIVERSITY OF IOWA.

* Neyman and Pearson, *loc. cit.*, p. 210; Rider, *loc. cit.*, p. 141.

A CERTAIN TRANSFORMATION ON METRIC SPACES.

By G. T. Whyburn.

1. In a study of compact locally connected continua Mazurkiewicz † has defined what amounts to a transformation on compact metric spaces K which is effected by changing all distances $\rho(x, y)$ in K into distances $\rho^*(x, y)$ where $\rho^*(x, y)$ is the greatest lower bound of the aggregate of numbers $[\delta(C)]$, where C is any continuum in K containing $x + y$ and $\delta(C)$ is the diameter of C. Now if we alter this definition merely by allowing C to be *any connected set* in K containing $x + y$, it is clear that the transformation will not be altered as applied to compact spaces K. However, in this form it is applicable to spaces which are not compact; and in particular, when applied to metric spaces M which are connected and locally connected, it turns out that, even though in this case the minimum diameter $\rho^*(x, y)$ is not necessarily effectively attained in the sense that there actually exist a connected set C such that $\delta(C) = \rho^*(x, y)$, nevertheless many important properties, including most of those found by Mazurkiewicz and Urysohn in the compact cases, are readily deducible. The present paper is devoted to the development of this transformation to such spaces M and to the application of the result found to the interesting case of a plane region.

It will be shown, among other results, that if we call this transformation T, then T is a homeomorphism whose inverse is uniformly continuous and which leaves diameters of connected sets invariant. Thus Property S is also invariant under T. Furthermore $T(M)$ is uniformly locally connected, and thus it follows that *any connected, locally connected metric space M is homeomorphic with a space M which is uniformly locally connected.*

In § 5 the results are applied to show that if R is any simply connected plane region having property S ‡ and with boundary B, then the boundary J of $T(R)$ in the space obtained by completing $T(R)$ after the manner of

† See *Fundamenta Mathematicae*, Vol. 1 (1920), pp. 167, 168. This same transformation was also studied later by Urysohn (see *Verhandelingen der Koninklyke Akademie van Wetenschappen te Amsterdam*, Erste Sectie, Deel XIII, No. 4, 1928, pp. 38-42) who also considered it only when applied to compact spaces.

‡ A set of points H has property S provided that for every $\epsilon > 0$, H is the sum of a finite number of connected sets each of diameter $< \epsilon$. (See Sierpinski, *Fundamenta Mathematicae*, Vol. 1, p. 44). It has been shown by R. L. Moore (*Fundamenta Mathematicae*, Vol. 3, p. 235) that a bounded plane simply connected region has property S if and only if its boundary is locally connected.

the Cantor addition of the irrational numbers to the rational number system, is a simple closed curve; and since the inverse of T is uniformly continuous, it can be extended in one and only one way to give a continuous transformation W of J into B. Furthermore, it will be shown that W is a *minimal mapping,* from the standpoint of multiplicity, of J onto B; *the multiplicity under W of any point p of B is equal to the number of components of $B - p$ if this number is finite and is infinite if this number is infinite.* Thus, under the transformation W, not only does the image of a point p on J traverse continuously the points of the boundary B in cyclic order, but the structure of B is accurately described by the transformation W.†

2. *The Transformation T.* We shall consider a connected and locally connected metric space which we denote by M. For each pair of points x and y of M let $\rho^*(x, y)$ be the greatest lower bound of the aggregate of diameters $[\delta(C)]$ of all connected sets C in M which contain both x and y. Now for each pair x, y in M, the number $\rho^*(x, y)$ is a finite non-negative real number. Obviously this is true if $\delta(M)$ is finite, which we may suppose without loss of generality. Even without assuming $\delta(M)$ finite, however, as Mr. C. H. Harry has remarked, we may still show that every $\rho^*(x, y)$ is finite by using the fact that x and y can be joined in M by a finite simple chain of connected regions each of diameter < 1 which is thus a connected set $C \supset x + y$ and with $\delta(C) \leq k$, where k is the number of links in the chain. Also, as will be shown below, the numbers $\rho^*(x, y)$ have distance character in that they satisfy the distance axioms necessary to define a metric space. Let M^* denote the space whose points are exactly the points of M but in which "distance" is defined by means of the above definition, i. e., the distance $\rho(x^*, y^*)$ between two points x^* and y^* of M^* is the number $\rho^*(x, y)$ determined in the above manner for the points x and y of M, where x and y are the points in M identified with the points x^* and y^* respectively in M^*. Thus M^* is the space obtained from M by changing all distances $\rho(x, y)$ into distances $\rho^*(x, y)$. If we regard this operation as a transformation T, then we have that $T(M) = M^*$. We denote points of M by $x, y, \cdots$ and corresponding points in M^* by $x^*, y^*, \cdots$. Distances in each space are denoted by the operator ρ. Thus we have always

(i) $$\rho^*(x, y) = \rho(x^*, y^*).$$

We proceed now to establish the following

† In this connection the reader is referred to a result relating the Menger-Urysohn order of a point of a regular curve and its multiplicity under a certain transformation of a circle into that curve, announced by G. Nöbeling in the *Wiener Akademie Anzeiger*, July, 1929, No. 17.

3. *Properties of the transformation.*†

(1) $\rho^*(x, y) \geqq \rho(x, y)$;

(2) *diameters of connected sets are invariant under* T, i. e., $\delta(C) = \delta(C^*)$, *where* C *is any connected set;*

(3) *the numbers* $\rho^*(x, y)$ *have distance character*, i. e., $\rho^*(x, y) > 0$, $x \neq y$; $\rho^*(x, x) = 0$; $\rho^*(x, y) = \rho^*(y, x)$; $\rho^*(x, z) \leqq \rho^*(x, y) + \rho^*(y, z)$; *thus* M^* *is a metric space;*

(4) T *is a topological transformation. Thus* M^* *is homeomorphic with* M*;*

(5) *property* S *is invariant under* T*;*

(6) M^* *is uniformly locally connected;*

(7) *the inverse of* T *is uniformly continuous;*

(8) *if* M *is uniformly locally connected, then* T *is uniformly continuous and conversely;*

(9) *all spherical neighborhoods in* M^* *are connected*, i. e., *for each point* p^* *of* M^* *and each* $r > 0$, *the set* $V_r(p^*)$ *of all points* x^* *in* M^* *with* $\rho(x^*, p^*) < r$ *is connected.*

Proof. If $C \supset x + y$, then $\delta(C) \geqq \rho(x, y)$. Therefore, g. l. b. $[\delta(C)] = \rho^*(x, y) \geqq \rho(x, y)$, where "g. l. b." means "greatest lower bound," and (1) is proved.

By (1) we have $\delta(C) \leqq \delta(C^*)$. Since $x + y \subset C$ implies $\rho^*(x, y) \leqq \delta(C)$, therefore we have $\delta(C^*) =$ l. u. b. $[\rho^*(x, y)] \leqq \delta(C)$, where x and y are in C and "l. u. b." means "least upper bound." These two inequalities give $\delta(C) = \delta(C^*)$, which is (2).

The first two relations in (3) are corollaries to (1) and (2) respectively. The third results from the fact that $x + y \subset C$ is the same relation as $y + x \subset C$. The fourth follows from the fact that if, in general, C_{ab} denotes a connected set containing $a + b$, then any set $C_{xy} + C_{yz}$ is a set C_{xz}. For then since $\delta\,[C_{xy} + C_{yz}] \leqq \delta(C_{xy}) + \delta(C_{yz})$, we have

$$\text{g. l. b. } [\delta(C_{xz})] \leqq \text{g. l. b. } [\delta(C_{xy})] + \text{g. l. b. } [\delta(C_{yz})],$$

which is the same as $\rho^*(x, z) \leqq \rho^*(x, y) + \rho^*(y, z)$.

That the transformation T is $(1-1)$ follows from (3). By virtue of the local connectivity of M it follows that when $\rho(x_n, x) \to 0$, also $\rho^*(x_n, x) = \rho(x^*_n, x^*) \to 0$; and by (1) we have that $\rho(x_n, x) \to 0$ when $\rho^*(x_n, x) \to 0$.

† For the case of compact spaces, properties (1) and (3) were established by Mazurkiewicz (*loc. cit.*), and property (4) was established by Urysohn (*loc. cit.*).

Therefore T and T^{-1} are continuous and thus T is topological, which proves (4).

If M has property S, then for any $\epsilon > 0$, $M = \sum_1^n C_i$, where each C_i is connected and $\delta(C_i) < \epsilon$. Then since by (2), $\delta(C^*_i) = \delta(C_i)$, we have $M^* = \sum_1^n C^*_i$, where $\delta(C^*_i) < \epsilon$. Thus M^* has property S, and (5) is proved.

To prove (6), let ϵ be any positive number. Take a number $\delta = \epsilon/2$. Then if x^* and y^* are any two points of M^* with $\rho(x^*, y^*) = \rho^*(x, y) < \delta$, by the definition of $\rho^*(x, y)$ there exists a connected set C in M containing $x + y$ and with $\delta(C) < 2\delta = \epsilon$. Then $C^* \supset x^* + y^*$, and, by (2), $\delta(C^*) = \delta(C) < \epsilon$. Thus M^* is uniformly locally connected.

Property (7) follows immediately from (1). For if ϵ is any positive number, then taking $\delta = \epsilon$ we have by (1) that if x^* and y^* are points in M^* with $\rho(x^*, y^*) = \rho^*(x, y) < \delta = \epsilon$, then $\rho(x, y) \leq \rho^*(x, y) < \epsilon$.

If M is uniformly locally connected, then for any $\epsilon > 0$, a $\delta > 0$ exists such that when $\rho(x, y) < \delta$, there exists a connected set C in M with $x + y \subset C$ and $\delta(C) < \epsilon$; thus it follows that $\rho(x, y) < \delta$ implies $\rho^*(x, y) = \rho(x^*, y^*) < \epsilon$, which proves that T is uniformly continuous. If, conversely, T is uniformly continuous, then for any $\epsilon > 0$, a $\delta > 0$ exists such that $\rho(x, y) < \delta$ implies $\rho(x^*, y^*) = \rho^*(x, y) < \epsilon$; therefore, by definition of $\rho^*(x, y)$, there exists a connected set C in M with $x + y \subset C$ and $\delta(C) < \epsilon$, which proves M uniformly locally connected. Thus (8) is established.

Finally, (9) results from the fact that $\rho(x^*, p^*) < r$ implies the existence of a connected set C in M with $x + p \subset C$ and $\delta(C) < r$; this gives $C \subset V_r(p^*)$, from which relation the connectivity of $V_r(p^*)$ is manifest.

4. *Extension of the domain of definition of T.* It has been shown by Hausdorff † that any metric space S may be "completed" into a complete space (i. e., a metric space in which every fundamental sequence of points has a limit point) by adding points on to S corresponding to every fundamental sequence in S which has no limit point in S in a manner entirely parallel to the Cantor method of defining the irrational numbers. Suppose we "complete" our spaces M and M^* in this way and call the resulting complete spaces M_o and M^*_o respectively. We shall consider here three cases.

(a). *In general* since every point of $M^*_o - M^*$ is a limit point of M^* and since, by property (7), T^{-1} is uniformly continuous, it follows by a well known theorem that the definition of T^{-1} can be extended to include the

† See *Grundzüge der Mengenlehre*, 1914, p. 315.

points of $M^*_c - M^*$ in one and only one way so that if we call the resulting transformation W, then W will be single valued and continuous and will be equal to T^{-1} on M^*. The image under W of any point p of $M^*_c - M$ must be a point of $M_c - M$; for if $W(p) = x$ where $x \subset M$, then $x^* \subset M^*$ and hence $x^* \neq p$; but then M^* contains a subset N^* having p but not x as a limit point; hence, W being continuous, the subset $N = W(N^*)$ of M has x as a limit point, contrary to the fact that M and M^* are homeomorphic. Thus we have proved that

$$\text{(ii)} \qquad W(M^*_c - M^*) \subset M_c - M.$$

(b) *In case M has property S*, then by property (6), M^* also has property S. Then since in the spaces M_c and M^*_c respectively, $\bar{M} = M_c$ and $\bar{M}^* = M^*_c$ it follows † that M_c and M^*_c have property S. Thus M_c and M^*_c are complete spaces having property S, and hence by a theorem of Hausdorff's,‡ they must be compact. Now in this case it follows that every point of $M_c - M$ must be the image under W of some point of $M^*_c - M^*$. For let p be any point of $M_c - M$. Then since, as is at once clear, $M + p$ is locally connected at p, there exists in M a sequence of connected sets $U_1 \supset U_2 \supset U_3 \supset \cdots$ such that $\delta(U_i) \to 0$ and for each n, $p \subset \bar{U}_n$. For each i, choose a point x_i in U_i. Then since, for each m, $\sum_m^\infty x_i \subset U_i$, it follows that $x^*_1, x^*_2, \cdots$ is a fundamental sequence in M^*. Hence this sequence has a limit point q in M^*_c. The point q cannot be in M^*, for the sequence $x_1, x_2, \cdots$ has as its limit point the point p of $M_c - M$. Thus $q \subset M^*_c - M^*$, and as W is continuous, $W(q) = p$. Thus we have shown that $W(M^*_c - M^*) \supset M_c - M$. By virtue of (ii) we have, in case M has property S, that

$$\text{(iii)} \qquad W(M^*_c - M^*) = M_c - M.$$

Therefore W is a single valued and continuous transformation mapping M^*

† See R. L. Moore, *loc. cit.*

‡ According to Hausdorff, a space is totally bounded (total beschränkt) provided it is the sum of a finite number of arbitrarily small sets. Hausdorff's theorem which we apply here states that any complete, totally bounded space is compact. See *Grundzüge der Mengenlehre*, 1914, p. 314.

In this connection we note the fact that since clearly the definition of T as well as properties (1), (2), (3), (5), (6), (7) have meaning and are valid whether M is locally connected or not (so long as we suppose, for convenience, that $\delta(M)$ is finite), it follows that *either one of the following conditions is both necessary and sufficient in order that a connected metric space M have property S:* (α) *that the space M^* be totally bounded or* (β) *that the space M^*_c be compact.* In case M is compact, property S is equivalent to local connectivity, $M^* = M^*_c$, and this proposition reduces to a theorem of Urysohn (*loc. cit.*).

onto M and $M^*_o - M^*$ onto $M_o - M$. Incidentally we have proved the following theorem:

If M is any connected set having property S, then M is homeomorphic with a uniformly locally connected subset M^* of a compact-locally connected continuum M^*_o.

(c). *If M is uniformly locally connected,* then by virtue of (7) and (8) T and T^{-1} are uniformly continuous. Consequently, T and T^{-1} can be extended to the points of $M_o - M$ and $M^*_o - M^*$, respectively, in a unique manner so that the resulting transformations are single valued and continuous on M_o and M^*_o respectively and so that their values on M and M^* are unaltered. Also it is seen at once that under the extended transformations T and T^{-1} it is true that for each point x in M_o (whether x be in M or not), $T^{-1}[T(x)] = x$, so that T^{-1} is the true inverse of T over the completed spaces M_o and M^*_o. Therefore, the extended transformation T is a topological transformation throwing M_o into M^*_o, M into M^*, and $M_o - M$ into $M^*_o - M^*$. Thus we have shown that M_o and M^*_o are homeomorphic and also that $M_o - M$ and $M^*_o - M^*$ are homeomorphic.

5. *Application to plane regions.* Let R be a plane connected region which has property S and which has a connected boundary B. Then R is a space M and, in the notation of this paper, we have at once that $R_o - R = B$. Let J denote the set of points $R^*_o - R^*$, where $R^* = T(R)$. Then, applying § 4, (a) and (b), we have a continuous transformation W of R^*_o into R_o such that $W = T^{-1}$ on R^* and $W(J) = B$. We now proceed to show that

(I) *The set J is a simple closed curve.*

First, J is compact and closed. For by § 4, (b), R^*_o is compact; and if J were not closed, some point x^* of R^* would be a limit point of J. But then as W is continuous, the point x of R would be a limit point of $W(J)$, contrary to the fact that $W(J) = B$ and no point of R is a limit point of B.

Second, J is connected. For if not, then with the aid of the Borel Theorem it is seen that there exists a closed (in R^*_o) and compact subset F^* of R^* such that $R^* - F^* = A^* + G^*$, where A^* and G^* are mutually separated and $\bar{A}^* \cdot J \neq 0 \neq \bar{B}^* \cdot J$ (the bar indicates closure in R^*_o). Then $F \subset R$, $R - F = A + G$, and A and G are mutually separated and $\bar{A} \cdot B \neq 0 \neq \bar{G} \cdot B$. A simple application of the Borel Theorem and of the arcwise connectivity property of R shows the existence of a continuum H in R which contains F in its interior, i. e., every point of F is an interior point of H. Let K be the boundary of the complementary domain D of H

which contains B. By a well known theorem, K is connected. Now since B contains limit points of each of the sets A and G, it follows that D contains points x and y of A and G respectively. But then if xu and yv are arcs in R with $xu \cdot H = u$ and $yv \cdot H = v$, the set $xu + K + yv$ is a connected subset of $R - F$ containing both x and y, contrary to the fact that F separates x and y in R. Therefore J is connected.

Finally, J is disconnected by the omission of any two of its points. For let a and b be any two points of J. Since by (5), R^* has property S, therefore † a and b are accessible from R^*. Accordingly it follows that there exists in R^*_0 an arc ab such that $(ab)^* = ab - (a + b) \subset R^*$. Then $W(ab)$ is either an arc or a simple closed curve according as $W(a)$ and $W(b)$ are different or the same and $(ab) = W(ab) - W(a) - W(b) \subset R$. In either case, $R - (ab) = D_1 + D_2$, where D_1 and D_2 are mutually separated non-vacuous sets. Therefore $R^* - (ab)^* = D^*_1 + D^*_2$ and D^*_1 and D^*_2 are mutually separated. Let $J_1 = \bar{D}^*_1 \cdot J - (a + b)$ and $J_2 = \bar{D}^*_2 \cdot J - (a + b)$. Then clearly $J_1 + J_2 = J - (a + b)$. Furthermore, J_1 and J_2 are non-vacuous, mutually separated sets. They are non-vacuous, for there exist points p_1 and p_2 (not necessarily distinct) in $B - W(a) - W(b)$ which are limit points of D_1 and D_2 respectively; D_1 and D_2 contain sequences of points N_1 and N_2 respectively which converge to p_1 and p_2 respectively; since R^*_0 is compact, N^*_1 and N^*_2 each have at least one limit point q_1 and q_2 respectively. Then $q_1 + q_2 \subset J$, and since $W(q_1) = p_1$ and $W(q_2) = p_2$, we have $q_1 + q_2 \subset J - (a + b)$; and since q_1 and q_2 are limit points of D^*_1 and D^*_2 respectively, we have $q_1 \subset J_1$ and $q_2 \subset J_2$, which proves J_1 and J_2 non-vacuous. Now if J_1 and J_2 are not mutually separated, it follows that there exists a point p in $J - (a + b)$ which is a limit point both of D^*_1 and of D^*_2. Let U be a neighborhood of p with $U \cdot ab = 0$. Then since, by (b), R^* is uniformly locally connected, it follows that there exist points x_1 and x_2 in D^*_1 and D^*_2 respectively which can be joined by a connected subset of $U \cdot R^*$. But this is impossible because every connected set in R joining D^*_1 and D^*_2 must contain a point of ab. Therefore J_1 and J_2 are mutually separated, and hence J is disconnected by the omission of any two of its points.

Thus we have shown that J is a compact continuum which is disconnected by the omission of any two of its points. It follows by a theorem of R. L. Moore ‡ that J is a simple closed curve.

Since $B = W(J)$ and W is continuous, it follows that B must be locally

† See G. T. Whyburn, *Proceedings of the National Academy of Sciences*, Vol. 13 (1927), p. 650 and *Fundamenta Mathematicae*, Vol. 12, p. 272.

‡ See *Transactions of the American Mathematical Society*, Vol. 21 (1920), p. 342.

connected. Hence as a corollary to (I) we have the result of R. L. Moore's † which states that *The boundary of every plane, connected, and simply connected region which has property S is a locally connected continuum.* Also, in case R is uniformly locally connected and bounded, it will have property S and by § 4, (c), it follows that B and J are homeomorphic. Thus in this case B is a simple closed curve, and as a second corollary to (I) we have R. L. Moore's theorem ‡ to the effect that *The boundary of every bounded, plane, connected, and uniformly locally connected region is a simple closed curve.*

Now let p be any point of B. Then since $W^{-1}(p)$ must cut J into at least as many components as there are components of $B - p$, it follows that the multiplicity $m(p)$ of p under W, i. e., the number of points x on J such that $W(x) = p$, is at least as great as the number $\alpha(p)$ of components of $B - p$. We shall now prove

(II) *If either of the numbers $\alpha(p)$ and $m(p)$ is finite, the other is finite and the two are equal.*

To this end, let $p_1, p_2, p_3, \cdots p_k$ be any set of k-points on J such that, for each $i \leq k$, $W(p_i) = p$. There exist arcs $x^*_i p_i$ such that $x^*_i p_i - p_i \subset R^*$ and such that for each j and $r \leq k$, $\rho\,[(x^*_j, p_j),\ (x^*_r, p_r)] > d$, where d is some fixed positive number independent of j and r. Then $[W(x^*_i, p_i)]$ are arcs $[x_i p]$ such that $x_i p - p \subset R$ and such that no connected subset of R of diameter $< d$ contains points of two of these arcs. Let C be a circle with center p which is of diameter $< d$ and which does not enclose any of the points $[x_i]$. For each i, the arc $x_i p$ contains an arc $z_i p$ such that $z_i \subset C$ and $z_i p - z_i \subset I$, where I is the interior of C. The arcs $z_1 p, z_2 p, \cdots z_k p$ divide I into exactly k regions $D_1, D_2, \cdots, D_k$. It follows at once that for each i, p is a limit point of $D_i \cdot B$ (for no two of the arcs $z_i p$ can be joined in $I \cdot R$) and, since B is locally connected, there exists in B an arc $u_i p$ such $u_i p - p \subset D_i$. Now no two of the sets $u_i p - p$ can lie together in the same component of $B - p$. For if say $u_i p - p$ and $u_j p - p$ lie together in a component N of $B - p$, then N contains an arc $u_i u_j$ and $u_i u_j + u_i p + u_j p$ contains a simple closed curve L which contains segments $v_i p$ and $v_j p$ of $u_i p$ and $u_j p$ respectively. But then since $v_i p - p \subset D_i$, it follows that both the interior and the exterior of L contain points of R, contrary to the fact that R is connected and $L \cdot R = 0$. Therefore no two of the sets $u_i p - p$ lie together in

† See R. L. Moore, *Fundamenta Mathematicae*, Vol. 3, p. 235.

‡ See R. L. Moore, *Proceedings of the National Academy of Sciences*, Vol. 4 (1918), pp. 364-370.

a single component of $B - p$; and since there are k of the sets $u_i p - p$, it follows that there are at least k components of $B - p$. Thus we have shown that if $W^{-1}(p)$ contains at least k points, $B - p$ has at least k components. Hence $\alpha(p) \geqq m(p)$, and if either of these numbers is finite the other is finite and the two are equal.

Finally we prove

(III) *W is not constant on any arc of J.*

To this end, let p be any point of B and let a and b be any two points of J (if two such points exist) such that $W(a) = p = W(b)$. As shown above, there exists an arc ab such that $(ab)^* = ab - (a + b) \subset R^*$. Then $W(ab)$ is a simple closed curve with $W(ab) - p = (ab) \subset R$ and $R - (ab) = D_1 + D_2$, where D_1 and D_2 are mutually separated and D_1 lies within $W(ab)$ and D_2 without $W(ab)$. Let x^* and y^* be points on $(ab)^*$. Then since $x \to p$ and $y \to p$ as $x^* \to a$ and $y^* \to b$ respectively, whereas $\lim\limits_{x,y \to p} \rho^*(x, y) = \rho(a, b)$, it follows that both the interior and the exterior of $W(a, b)$ must contain points q_1 and q_2 respectively of B. Since q_1 is not a limit point of D_2 and q_2 is not a limit point of D_1, it follows that if v_1 and v_2 are points on J such that $W(v_1) = q_1$ and $W(v_2) = q_2$, then v_1 must belong to one of the arcs of J from a to b and v_2 to the other one. Hence W is not constant on either of the arcs of J from a to b, for $W(a) = W(b) = p$, whereas $W(v_1) = q_1$ and $W(v_2) = q_2$ and $q_1 \neq p \neq q_2$; and from this it follows that W is not constant on any arc of J, because a and b were any two points on J having the same image point in B.

By way of recapitulation we note that since $R^* + J$ is homeomorphic with the unit circle plus its interior, the proofs given in this section establish the following

THEOREM. *If R is any plane connected region which has property S and has a connected boundary B, and if C denotes the unit circle and I its interior, then there exists a topological transformation W of I into R which can be extended to the circle C in such a way as to give a single valued and continuous transformation of C into B which is not constant on any arc of C and is such that the multiplicity* † *under this extended transformation of*

† It seems very probable to the author that the correspondence set up by Carathéodory (See *Mathematische Annalen*, Vol. 73, pp. 323-370) between the prime ends of the boundary of any bounded simply connected region and the points on a circle may, in the case we are here considering, have the same multiplicity relations as we have established for our transformation W.

any point p of B is equal to the number of components of $B-p$ when this number is finite and is infinite when this number is infinite.

In conclusion we remark that, in so far as the set B itself is concerned, we may obtain a mapping W of the circle C onto B having all the properties stated in the above theorem *where B is any compact locally connected continuum containing at least one simple closed curve and every maximal cyclic curve of which is a simple closed curve.* For by a theorem of W. L. Ayres,† any such continuum B, whether it lie in the plane or not, is homeomorphic with the boundary B' of a plane bounded region R. By a theorem of R. L. Moore's ‡ this region R has property S. Hence we can apply our theorem and obtain all the properties for B' and thus also for B, since B and B' are homeomorphic.

THE JOHNS HOPKINS UNIVERSITY.

† See Ayres, *Fundamenta Mathematicae*, Vol. 14, p. 92.
‡ See Moore, *Fundamenta Mathematicae*, Vol. 3, p. 232.

ON UPPER SEMI-CONTINUOUS DECOMPOSITIONS OF COMPACT CONTINUA.*

By W. A. Wilson.

1. *Introduction.* The complexities which may occur in the structure of continua have naturally led to various methods of analyzing them. One of the most natural is to decompose the given continuum into sub-continua as elements, which may be taken as "points" of a new locally connected continuum. The use of prime parts for this purpose by H. Hahn and R. L. Moore is well known.

It was with the idea of finding a finer decomposition that the writer invented a notion called the *oscillatory set* (See § 9) of a continuum about a point,† a notion which did not come up to expectations in general on account of a lack of uniqueness. However, in the case of a bounded continuum irreducible between two points it was found that the oscillatory set is unique and that, if no indecomposable continua except continua of condensation are present, the continuum can be exhibited as a simple arc having certain oscillatory sets called *complete* as the "points." The problem for the irreducible continuum was finally completely solved by C. Kuratowski,‡ who divided the continuum into elements called *tranches,* which are identical with the complete oscillatory sets of the writer for the case mentioned above and are connected sets of indecomposable continua and continua of condensation in the more general case.

It is the principal business of the present article to study this same problem for another class of continua for which the oscillatory sets can be taken as unique, namely that of compact one-dimensional continua which have finite connectivity. The space used is metric and compact.

2. *Definitions.* A decomposition of a compact continuum M into disjoint sub-continua $\{X\}$ is called *regular* if (a) $M = \Sigma\,[X]$; (b) the decomposition is upper semi-continuous; and (c) M is locally connected about

* Presented to the Society, September, 1931.

† "On the Oscillation of a Continuum at a Point," *Transactions of the American Mathematical Society,* Vol. 27 (1925), pp. 429-440; "On the Structure of a Continuum, Limited and Irreducible Between Two Points," *American Journal of Mathematics,* Vol. 48 (1926), pp. 147-168.

‡ "Theorie des continus irreductibles entre deux points II," *Fundamenta Mathematicae,* Vol. 10, pp. 225-275.

each X. The sub-continua $\{X\}$ are called *elements* of M, or elements of the decomposition D of M. The decomposition $M = \Sigma [X]$ is upper semi-continuous if, for every $\epsilon > 0$, there is a $\delta > 0$ such that $X' \subset V_\epsilon(X)$ if $X' \cdot V_\delta(X) \neq 0$.* The continuum M is locally connected about X, if for every $\epsilon > 0$, there is a $\delta > 0$ and a sub-continuum C of M for which $V_\delta(X) \subset C \subset V_\epsilon(X)$.

If D and E are two decompositions of a continuum M, E is said to be *finer* than D if *every* element of E is a sub-set of some element of D and *some* element of E is a *proper* sub-set of an element of D. A sequence $\{D_i\}$ of decompositions of a continuum M is called *descending* if each D_i is finer than D_{i-1}.

If a regular decomposition D is such that there is no finer regular decomposition, we say that the decomposition D is *irreducible*. The next few sections are devoted to a chain of theorems showing that a compact continuum which has any regular decomposition has an irreducible one.

3. As a preliminary it should be noted that, if D and E are upper semi-continuous decompositions of the compact continuum M and E is finer than D, then the decomposition E generates an upper semi-continuous decomposition of at least one element X of the decomposition D.

On the other hand, it is not true that, if D and E are regular, the decomposition of the element X generated by E is necessarily regular. To see this, consider the four plane sets defined as follows:

P: $x = 0$, $(0 \leqq y \leqq 1)$;
Q: $x = 1/2^n$, $(n = 1, 2, \cdots)$; $0 \leqq y \leqq 1$;
R: $y = 0$, $1/2$, or 1; $(0 \leqq x \leqq 1)$;
S: $y = m/2^n$, $m < 2^n$ and integral; $(0 \leqq x \leqq 1/2^n$; $n = 2, 3, 4, \cdots)$.

Then $M = P + Q + R + S$ is a compact locally connected continuum. If we take $P + Q + R$ as one element and the remaining points of M as the other elements, we have a regular decomposition D. If we then take the points of M as elements, we have a regular decomposition E of M which is finer than D. The points of the continuum $P + Q + R$ are the elements into which E decomposes it, but $P + Q + R$ is not locally connected.

4. THEOREM. *Let $\{D_i\}$ be a descending sequence of decompositions of the compact continuum M. For any point x of M let X_i be the element containing x in the decomposition D_i and X be the divisor of the sequence*

* The notation $V_\epsilon(X)$ signifies the set of points of M whose distances from X are less than ϵ. Owing to the compactness of M there is no loss in generality in using the same δ for every X.

$\{X_i\}$. *Then* $M = \Sigma [X]$ *is a decomposition* D *of* M *into disjoint continua and* D *is finer than every* D_i.

Proof. By definition of X_i and X, each point x of M lies in one and only one X_i; hence $M = \Sigma [X]$ is disjoint.

Each X is the divisor of a monotone decreasing sequence of compact continua $\{X_i\}$. Hence $M = \Sigma [X]$ is a decomposition D of M into disjoint continua.

By construction D is at least as fine as every D_i. Suppose that $D = D_i$ for some i. Since D_{i+1} is finer than D_i, some $X_{i+1} \subset X_i$ and $X_{i+1} \neq X_i$. But $X \subset X_{i+1}$, and so $X \neq X_i$, a contradiction. Hence D is finer than every D_i.

This completes the proof. For brevity we shall say that D is the limit of the sequence of decompositions $\{D_i\}$.

5. THEOREM. *Let* $\{X_i\}$ *be a monotone decreasing sequence of sub-continua of a compact continuum* M, *let* M *be locally connected about each* X_i, *and* X *be the divisor of the sequence* $\{X_i\}$. *Then* M *is locally connected about* X.

Proof. Let $\epsilon > 0$ and $\eta = \epsilon/3$. Then for some i_0,

$$X_i \subset V_\eta(X), \qquad i > i_0. \tag{1}$$

On the other hand, for a fixed $i > i_0$, there is a $\delta > 0$ and a sub-continuum F of M such that

$$V_\delta(X_i) \subset F \subset V_\eta(X_i). \tag{2}$$

Since $X \subset X_i$, (1) and (2) show that $V_\delta(X) \subset F \subset V_\epsilon(X)$, which gives the theorem.

COROLLARY. *Let* M *be a compact continuum and* x *be a point of* M. *Then there is a sub-set* X *of* M *irreducible with respect to the properties of being a continuum containing* x *and having* M *locally connected about* X.

6. THEOREM. *Let* $\{D_i\}$ *be a descending sequence of upper semi-continuous decompositions of a compact continuum and* D *be its limit. Then* D *is an upper semi-continuous decomposition.*

Proof. Let A be any element of D and A_i be the element of D_i containing A. The sequence $\{A_i\}$ is monotone decreasing and A is its divisor. Take $\epsilon > 0$ and $\eta = \epsilon/3$. For i greater than some i_0,

$$A_i \subset V_\eta(A). \tag{1}$$

Fix i. Since D_i is upper semi-continuous, there is a $\delta > 0$ such that, if X_i is any element of D_i for which $X_i \cdot V_\delta(A_i) \neq 0$,

(2) $$X_i \subset V_\eta(A_i).$$

Now let X be an element of D such that $X \cdot V_\delta(A) \neq 0$ and let X_i be the element of D_i containing X. Then by (1) and (2),

$$X \subset X_i \subset V_\eta(A_i) \subset V_\epsilon(A).$$

Hence D is upper semi-continuous by definition.

7. THEOREM. *Let $\{D_i\}$ be a descending sequence of regular decompositions of a compact continuum and D be its limit. Then D is a regular decomposition.*

This is an immediate consequence of §§ 4, 5, and 6.

8. THEOREM. *Let M be a compact continuum. Then there is an irreducible regular decomposition of M, if there is any regular decomposition.*

Proof. Consider the hyper-space H whose elements are the sub-continua of M with the distance between any two, A and B, defined as the lower bound of the numbers $\{r\}$ such that $A \subset V_r(B)$ and $B \subset V_r(A)$. It is known that, when M is compact, the space H is also compact.

With each regular decomposition D of M, we associate a set H_D in H, whose "points" are the elements $\{X\}$ of D and all the sub-continua of all the elements $\{X\}$. To save words we call the set of "points" $\{X\}$ the frame* of H . Since the decomposition of M is upper semi-continuous, it is easily seen that H_D is closed. Conversely, a closed sub-set H_D of H, which contains a sub-set $\{X\}$ of "points" which in M are disjoint continua forming a regular decomposition of M and whose remaining "points" are in M all the sub-continua of all the continua $\{X\}$ will be called a D-set.

It is easily seen that, if H_D and H_E are D-sets and H_E is a proper part of H_D, then the regular decomposition E is finer than D; and conversely.

Now let $\{H_i\}$ be a decreasing sequence of D-sets and K be the divisor. For each i a point x of M lies on a continuum X_i which is a "point" of the frame of H_i. If X is the divisor of the sequence $\{X_i\}$, the set of "points" $\{X\}$ in H constitutes a frame for K by § 7. Obviously all the sub-continua of the elements $\{X\}$ are "points" of K. Conversely, if Y is a "point" of K, it is one of every H_i and so in M it is a sub-continuum of some one X_i

* Note that this set is not in general a continuum. For the decomposition of M is merely upper semi-continuous and the definition of distance used in defining H is different from that used in treating the space of elements obtained in an upper semi-continuous decomposition.

for each i; hence Y is a sub-continuum of the divisor X of $\{X_i\}$. Thus K is a D-set.

We see now that the property of being a D-set satisfies the conditions of Brouwer's induction theorem and so H contains an irreducible D-set. But then the corresponding regular decomposition of M is also irreducible.

Remarks. The limitations of the above theorem should not be overlooked. Although it has been shown that for a compact continuum there is an irreducible regular decomposition, it has not been shown that there is only one, nor is there any indication as to how the elements are to be found. As far as the writer knows, this problem has been solved only for special cases,—for example, the irreducible continuum. We now turn to another large class of continua for which a solution is at hand.

9. *Oscillatory sets.* In the first article referred to in § 1 an oscillatory set of a compact continuum M about one of its points x was defined as follows. Let $\{\delta_i\}$ be a decreasing sequence of numbers converging to zero. For each i let V_i be the set of points of M whose distances from x are less than δ_i. Let X_1 be a sub-continuum of M irreducible about V_1, X_2 a sub-continuum of X_1 irreducible about V_2, etc. Then the divisor X of the sequence $\{X_i\}$ is an oscillatory set of M about X. Obviously X is a continuum.

In general this is not unique, but this difficulty can be avoided for the class of one-dimensional compact continua of finite connectivity. Such continua we call *m cyclic* *; if $m = 0$, we call them *acyclic*. The principal thing which makes these continua easy to handle, as will appear in the demonstrations, is the fact that the divisor of any two sub-continua of an m cyclic continuum has at most $m + 1$ components and therefore a relative inner point of both sub-continua is a relative inner point of a component of their divisor.

THEOREM. *Let M be an acyclic one-dimensional compact continuum. Then M has a unique oscillatory set about each of its points $\{x\}$.*

Proof. For each $\delta > 0$ there is at least one sub-continuum X_δ of M irreducible about $V_\delta(x)$. Suppose that there were another, Y_δ. Then $X_\delta \cdot Y_\delta$ is not a continuum, because $V_\delta(x) \subset X_\delta \cdot Y_\delta \subset X_\delta$.

But then $X_\delta \cdot Y_\delta$ has more than one component, which is a contradiction by the above remarks. Thus there is a unique sub-continuum of M irreducible about each $V_\delta(x)$, and consequently the oscillatory set X is unique.

---

* The number m is one less than P. Alexandroff's connectivity number. See his paper, "Über kombinatorische Eigenschaften allgemeiner Kurven," *Mathematische Annalen*, Vol. 96, pp. 512-554.

10. It seems probable that, for a compact continuum of finite cyclic number and a given point x, there is a unique sub-continuum irreducible about $V_\delta(x)$, if δ is small enough. This would insure the uniqueness of the oscillatory set. In the absence of such a proof we introduce the following notion. An oscillatory set of a continuum about a point x is called *irreducible* if no proper sub-set is an oscillatory set about x.

THEOREM. *Let M be an m cyclic compact one-dimensional continuum. If C and D are closed sub-sets, each of which contains an oscillatory set about the point x, then $C \cdot D$ has the same property.*

Proof. It is enough to prove this for the case that C and D are oscillatory sets and neither contains the other. Then there are two descending sequences of vicinities $\{V_i\}$ and $\{W_i\}$, each converging to x, and two monotone decreasing sequences of continua $\{C_i\}$ and $\{D_i\}$, irreducible about the respective vicinities $\{V_i\}$ and $\{W_i\}$ and converging to C and D respectively, and for no i does C_i contain D_i or D_i contain C_i.

Let K_1 be the component of $C_1 \cdot D_1$ containing x. Since $C_1 \cdot D_1$ has a finite set of components and contains either V_1 or W_1, there is an i_2 for which $K_1 \supset V_{i_2} + W_{i_2}$. Then $K_1 \cdot C_{i_2} \cdot D_{i_2}$ contains the smaller of the vicinities V_{i_2} and W_{i_2} and hence K_1 contains a continuum E_2 irreducible about the smaller vicinity. Clearly $E_2 \subset C_1 \cdot D_1$.

Let K_2 be the component of $E_2 \cdot C_{i_2} \cdot D_{i_2}$ containing x; as above it contains $V_{i_3} + W_{i_3}$ for some i_3. Let E_3 be the sub-continuum of K_2 irreducible about the smaller of the vicinities V_{i_3} and W_{i_3}. Clearly $E_3 \subset C_{i_2} \cdot D_{i_2}$.

Continuing indefinitely, we define an oscillatory set E, which is the divisor of the sequence $\{E_n\}$ and is contained in $C \cdot D$.

11. THEOREM. *Let M be an m cyclic one-dimensional compact continuum. Then for each point x there is one and only one irreducible oscillatory set.*

Proof. To show the existence of an irreducible oscillatory set it is sufficient by the Brouwer induction theorem to show that, if $\{E_n\}$ is a descending sequence of closed sets of which each contains an oscillatory set about x and E is the divisor of this sequence, then E has the same property.

Let the decreasing sequence $\{\epsilon_i\}$ converge to zero. Then there is some E_n, which we call E_1, a vicinity V_1 of x, and a continuum C_1 chosen from those converging to some oscillatory set about x contained in E_1, such that C_1 is irreducible about V_1 and is contained in $V_{\epsilon_1}(E)$. Likewise there is an E_2 following E_1 in the sequence $\{E_n\}$, a vicinity V_2 of x of radius less than half

that of V_1, and a continuum C_2 chosen from those converging to some oscillatory set about x contained in E_2, such that C_2 is irreducible about V_2 and is contained in $V_{\epsilon_2}(E)$. Obviously this process can be carried on indefinitely.

Let $F_1 = C_1$. If $C_2 \subset C_1$, let $F_2 = C_2$. Clearly C_2 does not contain C_1 as a proper part, because $V_1 \supset V_2$ and C_2 is irreducible about V_2. If C_2 contains points not on C_1, let K_1 be the component of $C_1 \cdot C_2$ which contains x. Then there is a sub-continuum F_2 of K_1 irreducible about some V_i, which we call V_{i_2}. Then $V_{i_2} \subset F_2 \subset V_{\epsilon_2}(E)$.

If $F_2 \supset C_{i_2+1}$, let $F_3 = C_{i_2+1}$. If not, $F_2 \cdot C_{i_2+1}$ has a component K_2 which contains a sub-continuum irreducible about some V_{i_3}. Then $V_{i_3} \subset F_3 \subset V_{\epsilon_3}(E)$. Since every $i_k \geqq k$, a continuation of this gives an oscillatory set F which is the divisor of the monotone decreasing sequence $\{F_k\}$ and, since $\epsilon_k \rightarrow 0$, $F \subset E$.

Hence by the Brouwer theorem there is a set irreducible with respect to the property of containing, and therefore of being, an oscillatory set about x. There cannot be two such, as by the previous theorem their divisor contains an oscillatory set.

Definition. A continuum M will be called *regular* if it has a unique irreducible oscillatory set X about each point x and X is contained in every oscillatory set of M about x. The continuum M in the above theorem is clearly regular. Evidently X is the divisor of all the sub-continua of M containing x as a relative inner point.

12. THEOREM. *Let M be a regular compact continuum. Let the sequence $\{x_i\}$ of points of M converge to x, and $\{X_i\}$ and X be the respective irreducible oscillatory sets. Then X contains the upper closed limiting set of the sequence $\{X_i\}$.*

Proof. For any $\epsilon > 0$ there is a $\delta > 0$ and a sub-continuum X_δ for which $X + V_\delta(x) \subset X_\delta \subset V_\epsilon(X)$. For some i_0, x_i lies in $V_\delta(x)$ if $i > i_0$. Hence some $V_\eta(x_i) \subset V_\delta(x)$. Then the previous inclusions show that $X_i \subset V_\epsilon(X)$, which proves the theorem.

13. THEOREM. *Let M be a regular compact continuum, and let M be locally connected about the sub-continuum K. Then K contains the irreducible oscillatory set of M about each point of K.*

Proof. For each $\epsilon > 0$ there is a $\delta > 0$ and a sub-continuum K_δ for which $V_\delta(K) \subset K_\delta \subset V_\epsilon(K)$. If x lies in K, $V_\delta(x) \subset V_\delta(K)$; consequently for each δ the irreducible oscillatory set X about x is contained in K_δ. Hence $X \subset K$.

COROLLARY. *If M satisfies the above hypotheses and K is the irreducible*

oscillatory set about some point x, M is not locally connected about any proper sub-continuum of K containing x.

14. THEOREM. *Let M be a regular compact continuum, and let K be a sub-continuum containing the irreducible oscillatory set of M about each of its points. Then M is locally connected about K.*

Proof. Let x be any point of K and X be its irreducible oscillatory set. Then for any $\epsilon > 0$ there is a $\delta > 0$, depending on x, and a continuum X_δ such that $V_\delta(x) \subset X_\delta \subset V_\epsilon(X)$. By the Borel theorem there is a finite number of the sets $\{V_\delta(x)\}$ whose union covers K; let K_δ be the union of the corresponding continua $\{X_\delta\}$. Then for some $\eta > 0$, $V_\eta(K) \subset K_\delta \subset V_\epsilon(K)$. Thus M is locally connected about K.

15. THEOREM. *Let M be a regular compact continuum, and let $M = \Sigma [K]$ be a decomposition into disjoint continua, each of which is the irreducible oscillatory set about some of its points and contains the irreducible oscillatory sets about all of its points. Then the decomposition is a unique irreducible regular decomposition.*

Proof. By § 14, M is locally connected about each K. Let $\{K_i\}$ be a sequence of elements, each containing a point y_i, and let $y_i \to b$, a point of K'. Let x_i be a point about which K_i is the irreducible oscillatory set. For a sub-sequence $\{x_{i_n}\}$ $x_{i_n} \to a$. If A is the irreducible oscillatory set about a, A contains the upper closed limit of the sequence $\{K_{i_n}\}$, by § 12, and consequently $A \supset b$. Since b also lies in K', it follows that $A \subset K'$. As this is true for any convergent subsequence of the points $\{x_i\}$, we see that K' contains the upper closed limit of $\{K_i\}$ and so the decomposition is regular.

Now let $M = \Sigma [L]$ be any other regular decomposition whatever. Any element K is an irreducible oscillatory set about some definite point x, which lies in a definite L. Then $K \subset L$ by § 13. Hence the decomposition $M = \Sigma [K]$ is a finer decomposition than $M = \Sigma [L]$, and is therefore irreducible.

16. THEOREM. *Let M be a regular compact continuum. Then each point x of M lies on a set K_x saturated with respect to the property of being a sub-continuum of M which contains x, is the union of irreducible oscillatory sets, and has each pair of its points on a connected sub-set consisting of a finite or enumerable set of irreducible oscillatory sets.*

Proof. Since the irreducible oscillatory set X about x has the property above described, we have only to show the existence of a saturated set of this nature. As M is compact, every well-ordered increasing sequence of sub-continua is enumerable. Hence we need only to show that, if $\{K_i\}$ is an

increasing sequence of the sets in question, there is another such set containing the union of the sequence.*

Let U be the union of the sequence $\{K_i\}$. Clearly, if U is closed, it has the desired property. If U is not closed, let W be the union of U and all the irreducible oscillatory sets $\{X\}$, for which $X \cdot \bar{U} \neq 0$. Obviously W is connected. We now proceed to show that it is closed. Let $\{x_i\}$ be a sequence of points in W and $x_i \to a$. If an enumerable set of these points lies in U, a lies in W by the definition of W. Let us assume, then, that no x_i lies in U. Then each x_i lies in the irreducible oscillatory set Y_i of some point y_i and $\bar{U} \cdot Y_i \neq 0$. Let b_i be a point of $\bar{U} \cdot Y_i$. For some partial sequence $\{y_i\}$ and $\{b_i\}$ converge to points y and b, respectively. By § 12 the upper closed limit of the sequence $\{Y_i\}$ is contained in Y, the irreducible oscillatory set about y. Hence $a + b \subset Y$. But b lies in $\bar{U}$ and so $Y \subset W$. This shows that a lies in W, or that W is closed.

It remains to show that any two points a and b of W can be joined by an at most enumerable connected chain of irreducible oscillatory sets contained in W. This is obvious, if $a + b \subset U$. Suppose now that a lies in U and b in $W - U$. Then a lies in some K_i and b lies in an irreducible oscillatory set B containing a limiting point of U; that is, there is a sequence $\{a_i\}$ of points of U converging to a point of B. Now a and a_1 lie on a connected sub-set C_1 of U composed of an at most enumerable set of irreducible oscillatory sets, a_1 and a_2 lie on a similar set C_2, etc. If we let C be the union of the sets $\{C_i\}$, C is connected and, since $B \cdot \bar{C} \neq 0$, so is $B + C$. Then $B + C$ is the required connected set. Finally the case that both a and b lie in $W - U$ is readily deduced from this one. In the light of the first paragraph, this completes the proof.

17. The continua defined in the previous section will be called *irreducible elements*. It is evident that: (a) each point x lies on one and only one element K_x; (b) the decomposition of M into irreducible elements is disjoint; (c) each element contains the irreducible oscillatory set of each of its points.

In consequence of § 11, an m cyclic compact one-dimensional continuum M is either itself an irreducible element or can be decomposed uniquely into irreducible elements. In the latter case it may be called *separable*.

18. THEOREM. *Let M be an m cyclic one-dimensional compact continuum and be separable. Then the decomposition of M into irreducible elements is an irreducible regular decomposition and is the only one possible.*

---

* See F. Hausdorff, *Mengenlehre*, pp. 173, 174.

Proof. Let $M = \Sigma [K]$ be the decomposition in question. This is locally connected by § 14. Let $\{K_i\}$ be a sequence of the elements, a_i be a point in K_i, and the sequence $\{a_i\}$ converge to a point a in some element K. If K does not contain the upper closed limit of the sequence $\{K_i\}$, there is a sub-sequence $\{K_{i_n}\}$ of elements, which we call $\{K_n\}$ to simplify the notation, each containing a point b_n, such that the sequence $\{b_n\}$ converges to a point b in some element L differing from K.

Since M is locally connected about each element, there is for any positive ϵ, which we may take less than one-third the distance between K and L, a $\delta > 0$ and continua K_δ and L_δ such that

$$V_\delta(K) \subset K_\delta \subset V_\epsilon(K) \quad \text{and} \quad V_\delta(L) \subset L_\delta \subset V_\epsilon(L).$$

Obviously $K_\delta \cdot L_\delta = 0$. On the other hand there is an n_0 such that, if $n > n_0$, $K_n \cdot K_\delta \neq 0$ and $K_n \cdot L_\delta \neq 0$. Taking $r > m + 1$, we see that $K_\delta + \Sigma [K_n]$ and $L_\delta + \Sigma [K_n]$, $n = n_0 + 1, n_0 + 2, \cdots, n_0 + r$, are two sub-continua of M whose divisor has more than $m + 1$ components, a contradiction since M is m cyclic. Hence the assumption that K does not contain the upper closed limit of $\{K_i\}$ is false, and the decomposition is upper semi-continuous. It is then regular.

It remains to show that it is finer than any other regular decomposition $M = \Sigma [L]$. Let a and b be points of some K lying on different elements of the second decomposition. By the definition of irreducible element, a and b lie on a connected sub-set C of K, which is the union of an at most enumerable set of irreducible oscillatory sets $\{X_i\}$. By § 13 each of these oscillatory sets lies in one and only one L, which we may call L_i. Then the union of the sets $\{L_i\}$ is connected. This is impossible, since they are not all identical and since the sets $\{L\}$ are mutually disjoint.

Hence every K is a sub-set of some L. This proves the theorem.

19. As seen by an example in the second article referred to in § 1, the irreducible regular decomposition is in general effectively finer than that into prime parts. The continuum of irreducible elements obtained is an at most m cyclic continuous curve, and the notions of end-points, cut-points, etc., can be immediately generalized. A more detailed study requires an investigation of the nature of the irreducible elements themselves.

Finally it may be remarked that the conclusions of §§ 17 and 18 are also valid for a continuum irreducible between two points, whether or not one-dimensional or m cyclic. Since Kuratowski's decomposition of such continua into *tranches* was an irreducible regular decomposition, it follows that for such continua the *tranche* and the irreducible element are identical.

YALE UNIVERSITY,
NEW HAVEN, CONN.

A COMPLETE CHARACTERIZATION OF PROPER PSEUDO *D*-CYCLIC SETS OF POINTS.

By Leonard M. Blumenthal.

1. A set S of undefined elements, for suggestiveness called "points" is called a *semi-metrical space* provided that for each two elements p, q of S there corresponds a not negative real number, called the "distance" between the points p and q, such that, denoting this number by pq, we have $pq = qp$, while $pq = 0$ if and only if the points p and q are identical. If for two pairs of points p, q; p', q' we have the relation $pq = p'q'$ we shall say that the two pairs of points are *congruent*. A mapping of a set S upon a set S' is called a congruent mapping if to each pair of points of S there corresponds a congruent pair of points of S'. Finally, two sets S and S' are called congruent provided there exists a congruent mapping of one upon the other.

Karl Menger has characterized the n-dimensional euclidean space R_n among general semi-metrical spaces by means of relations between the distances of its points.* He has shown that the R_n has the congruence order $n+3$. This means that every semi-metrical space, each $n+3$ points of which is congruent with $n+3$ points of the R_n, is congruent with a subset of the R_n. It is further shown that each semi-metrical space containing *more than* $n+3$ points, each $n+2$ of which is congruent with $n+2$ points of the R_n, is congruent with a subset of the R_n. The notion embodied in this important result is expressed by saying that the R_n has the *quasi-congruence* order $n+2$. Thus, if a set of points is such that each $n+2$ of the points is congruent with $n+2$ points of the R_n, while the whole set is not congruent with a subset of the R_n, the set must consist of exactly $n+3$ points. Such sets are called pseudo-euclidean.

Let us denote by S_n the n-dimensional sphere; that is, S_0 is a pair of points, S_1 is a circle (the term "circle" denoting throughout this paper a curve), S_2 is the surface of a sphere is three-space, etc. It has been shown that the space S_n has the congruence order $n+3$ also; *but it has not the quasi-congruence order* $n+2$. Thus, for the circle S_1 the sets analogous to

* Menger, "New Foundations of Euclidean Geometry," *American Journal of Mathematics*, Vol. 53 (1931), pp. 721-745. In this paper, Menger has shortened and revised proofs that appear originally in his paper, "Untersuchungen über allegemeine Metrik," *Mathematische Annalen*, Vol. 100, p. 113.

the pseudo-euclidean sets for R_1, which we shall call pseudo d-cyclic sets, are not confined to sets of exactly four points.*

In this paper we plan to characterize completely those *pseudo d-cyclic sets of points that contain neither a convex tripod* nor a *pseudo-linear* quadruple; that is to say, no four of the points contained in the set are such that one point lies between each two of the three others and no non-linear quadruple of the set has all four of its triples linear.† Such pseudo d-cyclic sets will be called *proper*.‡

2. We shall take as our point of departure, the following lemmas, the proofs of which we omit:

LEMMA 1. *In order that three points p_1, p_2, p_3, form a d-cyclic triple (that is, be congruent to three points of a circle of length 2d) it is necessary and sufficient that no two of the points have a distance greater than d, and that either the points satisfy the relation $p_1p_2 + p_2p_3 + p_3p_1 = 2d$, or one of the points lies between the two others.*

LEMMA 2. *A proper pseudo d-cyclic quadruple does not contain two linear triples.*

This lemma is obtained by showing that a pseudo d-cyclic quadruple can not contain exactly two linear triples, and then showing that a pseudo d-cyclic quadruple contains exactly three linear triples if and only if the four points form a convex tripod. Further, a proper pseudo d-cyclic quadruple can not contain four linear triples; for in this case the quadruple is either linear or pseudo-linear. The quadruple cannot be pseudo-linear since, by hypothesis, it is proper; it cannot be linear, since it would then be d-cyclic and not pseudo d-cyclic. Lemma 2, then, follows.

LEMMA 3. *A d-cyclic quadruple has at least two linear triples.*

From the last two lemmas, we have the theorem:

THEOREM 1. *In order that four points (p_1, p_2, p_3, p_4) be d-cyclic, it is necessary and sufficient that each three of the points be d-cyclic, that the*

---

* The distance between two points of S_1 is measured by the length of the shorter arc joining the two points. The length of the circle is taken as $2d$.

† The point q is said to lie between the points p and r provided that $pq + qr = pr$. The three points are said to be *linear*, since an inspection of the Heronian formula for the area of a triangle shows that if one point lies between two others, the three points are congruent to three points of a line. Evidently, a convex tripod can not be imbedded in a circle.

‡ That a pseudo-linear quadruple, with all its triples d-cyclic, may be pseudo d-cyclic was communicated to the writer by Miss Laura Klaufer, of Vienna, who attributed the remark to Mr. Franz Alt.

points do not form a convex tripod, that at least two of the triples contained in the four points be linear; and that in case the four points are pseudo-linear we have $p_1p_3 = p_2p_4 = d$; $p_1p_2 + p_2p_3 = p_1p_3$.

We shall prove first the following lemma.

LEMMA. *A proper pseudo d-cyclic quadruple has none of its triples linear.*

Let the quadruple p_1, p_2, p_3, p_4 be a proper pseudo d-cyclic quadruple. The quadruple can not contain two linear triples according to Lemma 2. In order to show that it does not contain a single linear triple, we suppose that one triple, say p_1, p_2, p_3, is the only linear triple that it contains, and we deduce from this a contradiction.

Since all of the triples are d-cyclic, and the triple p_1, p_2, p_3 is linear, we have the relations

$$(1) \qquad (p_1p_2 + p_2p_3 - p_3p_1)(p_1p_2 - p_2p_3 + p_3p_1)(p_1p_2 - p_2p_3 - p_3p_1) = 0.$$

$$(2) \qquad p_1p_2 + p_2p_4 + p_4p_1 = 2d.$$

$$(3) \qquad p_2p_3 + p_3p_4 + p_4p_2 = 2d.$$

$$(4) \qquad p_1p_3 + p_3p_4 + p_4p_1 = 2d.$$

Now from (2), (3), and (4) we obtain

$$p_1p_2 + p_2p_3 - p_3p_1 = 2d - 2(p_2p_4),$$
$$p_1p_2 - p_2p_3 + p_3p_1 = 2d - 2(p_1p_4),$$
$$p_1p_2 - p_2p_3 - p_3p_1 = -2d + 2(p_3p_4).$$

Substituting in (1) we obtain

$$(d - p_2p_4)(d - p_1p_4)(d - p_3p_4) = 0,$$

whence one of the three distances p_2p_4, p_1p_4, p_3p_4, is equal to d. But if the distance between two points is equal to d, it is evident that any d-cyclic triple containing these two points is linear. Thus, in any one of the three cases above, some triple other than the triple p_1, p_2, p_3 is linear, and the contradiction sought is obtained.

COROLLARY. *If the quadruple* p_1, p_2, p_3, p_4 *does not form a convex tripod, nor a pseudo-linear set, and is such that each triple is d-cyclic, and one of the triples is linear, then another triple is linear, and the quadruple is d-cyclic.*

3. *Characterization of proper pseudo d-cyclic quadruples.* Let us consider the proper pseudo d-cyclic quadruple p_1, p_2, p_3, p_4. Since by the lemma, no one of the triples that the set contains is linear, but all of the triples are d-cyclic, we have the relations:

(1) $p_1p_2 + p_2p_3 + p_3p_1 = 2d$; (3) $p_2p_3 + p_3p_4 + p_4p_2 = 2d$;

(2) $p_1p_2 + p_2p_4 + p_4p_1 = 2d$; (4) $p_1p_3 + p_3p_4 + p_4p_1 = 2d$.

Adding the relations (1) and (2) and using (3) and (4) we obtain $p_1p_2 = p_3p_4$. In a similar way we get $p_2p_3 = p_4p_1$ and $p_1p_3 = p_2p_4$. Substitution of these four values in the above four relations reduces each of them to a single relation, $a + b + c = 2d$, where we have written $a = p_1p_2 = p_3p_4$, $b = p_2p_3 = p_4p_1$, $c = p_3p_1 = p_4p_2$.*

Thus, a proper pseudo d-cyclic quadruple is seen to have its opposite distances equal. It is important to observe, however, that no one of the distances a, b, c is the sum of the other two; for in this case, the quadruple would have all of its triples linear, which we have seen is impossible for a proper pseudo d-cyclic quadruple. It is interesting to compare this with the analogous situation in the case of the R_1. It is known that pseudo-linear quadruples (that is, quadruples not congruent to four points of a line, while each of the triples it contains is linear) have their opposite sides equal, and *one of the sides is the sum of the other two.*† Pseudo-linear quadruples, however, may be imbedded in a circle; *but not necessarily in a circle of metrical diameter d.*

We have, then, established the following theorem completely characterizing proper pseudo d-cyclic quadruples:

THEOREM. *A proper pseudo d-cyclic quadruple is characterized by the fact that its opposite distances are equal, that no one of these distances is equal to the sum of the other two, and that the sum of three of the distances, no two of which are opposite, is equal to the length of the circle.*

COROLLARY 1. *There is a double infinity of proper pseudo d-cyclic quadruples.*

COROLLARY 2. *A pseudo d-cyclic quadruple has no two points with a distance equal to d; i. e., no two points are diametral.*

4. *Characterization of proper pseudo d-cyclic quintuples.* The fundamental theorem for these quintuples is proved by the aid of the following two lemmas:

---

* These relations might be obtained also by considering the four equations in the six distances determined by the four points. The matrix of the coefficients is seen to be of rank four, and hence there is a double infinity of solutions of the set. The above relations are valid if the hypothesis is weakened to exclude only sets containing convex tripods.

† Menger, "Untersuchungen über allgemeine Metrik," *Mathematische Annalen*, Vol. 100 (1928), p. 127.

LEMMA 1. *A proper pseudo d-cyclic quintuple does not contain any diametrical points.**

Suppose that two of the points, say p_1 and p_5, contained in the proper pseudo d-cyclic quintuple p_1, p_2, p_3, p_4, p_5 are diametral. Then $p_1p_5 = d$, and we have the relations

$$(1) \qquad p_1p_2 + p_2p_5 = d; \qquad p_1p_4 + p_4p_5 = d; \qquad p_1p_3 + p_3p_5 = d.$$

Now, since the circle has the congruence order four, at least one of the quadruples contained in the five points is pseudo d-cyclic. By corollary 2 of section 3, this quadruple does not contain both p_1 and p_5. We may assume the labelling so that the quadruple p_2, p_3, p_4, p_5 is pseudo d-cyclic. Since the set of five points is, by hypothesis, a proper set, the pseudo d-cyclic quadruple p_2, p_3, p_4, p_5 is proper, and we may apply to it the fundamental theorem of section 3. We have

$$p_2p_3 = p_4p_5; \qquad p_2p_5 = p_3p_4; \qquad p_2p_4 = p_3p_5.$$

Writing $p_2p_3 = a$, $p_2p_5 = b$, $p_2p_4 = c$, we have $a + b + c = 2d$.

Substituting in the relations (1) we obtain $p_1p_2 = d - b$; $p_1p_3 = d - c$; $p_1p_4 = d - a$, and all of the ten distances determined by the five points are expressed in terms of a, b, and c.

Consider, now, the quadruple p_1, p_2, p_3, p_4. We observe that the point p_1 *lies between each pair of the other three points*; that is, the four points form a convex tripod. Then the five points p_1, p_2, p_3, p_4, p_5 are seen to be, contrary to the hypothesis, an improper pseudo d-cyclic set. Hence, the assumption that a pair of points contained in the five points is diametral is seen to lead to a contradiction, and the lemma is proved.†

LEMMA 2. *A proper pseudo d-cyclic quintuple does not contain a linear triple.*

To establish this lemma, we assume that some triple contained in the quintuple is linear, and we show that this assumption leads to a contradiction. Suppose, then, that the triple p_1, p_2, p_3 is linear.‡ Then the two quadruples containing this triple, namely, p_1, p_2, p_3, p_4 and p_1, p_2, p_3, p_5, are d-cyclic, by

---

* This lemma can be proved under the weaker hypothesis that the set contains no convex tripods.

† It is not true, however, that a pseudo d-cyclic quintuple that does not contain any diametral points is necessarily proper. In the latter part of this section, an example is given of an improper pseudo d-cyclic quintuple without diametral points.

‡ It is sufficient, of course, to obtain the contradiction in this case since the same proof will hold, *mutatis mutandis*, if some other triple should be chosen linear.

the corollary in section 2, and hence each contains a second linear triple. We may select p_1, p_2, p_4 in the first quadruple, and p_1, p_2, p_5 in the second quadruple to be linear. Then the quadruple p_1, p_2, p_4, p_5 is evidently d-cyclic.

This leaves two quadruples to be examined; namely, p_1, p_3, p_4, p_5 and p_2, p_3, p_4, p_5. We show now that one of these quadruples is d-cyclic; for, suppose that both quadruples are pseudo d-cyclic. Then, if we denote by p_i, p_j, p_k any one of the triples contained in these two quadruples, we have

$$\text{(a)} \qquad p_ip_j + p_jp_k + p_kp_i = 2d,$$

and using the results of section 3, we have the relations

$$\text{(b)} \qquad \begin{aligned} p_1p_3 &= p_4p_5 = p_2p_3, \\ p_1p_4 &= p_3p_5 = p_2p_4, \\ p_1p_5 &= p_3p_4 = p_2p_5. \end{aligned}$$

But since the three triples p_1, p_2, p_3; p_1, p_2, p_4; p_1, p_2, p_5 are linear, we have

$$\begin{aligned} (p_1p_2 + p_2p_3 - p_3p_1)(p_1p_2 - p_2p_3 + p_3p_1)(p_1p_2 - p_2p_3 - p_3p_1) &= 0, \\ (p_1p_2 + p_2p_4 - p_4p_1)(p_1p_2 - p_2p_4 + p_4p_1)(p_1p_2 - p_2p_4 - p_4p_1) &= 0, \\ (p_1p_2 + p_2p_5 - p_5p_1)(p_1p_2 - p_2p_5 + p_5p_1)(p_1p_2 - p_2p_5 - p_5p_1) &= 0. \end{aligned}$$

Upon substituting from the relations (b) the above three relations yield

$$p_1p_2 = 2(p_1p_3) = 2(p_1p_4) = 2(p_2p_4) = 2(p_2p_5) = 2(p_1p_5).$$

Thus, from (b), the distance p_1p_2 is equal to the double of any one of the other nine distances determined by the five points. Substituting in any one of the relations (a) we see that each of these nine distances is equal to $2d/3$, and hence p_1p_2 is equal to $4d/3$. But this is impossible, since the points p_1 and p_2 are congruent to two points of a circle of length $2d$ and hence their distance can not exceed d.

We have thus shown that one of the quadruples p_1, p_3, p_4, p_5; p_2, p_3, p_4, p_5, is d-cyclic (since the contrary assumption has been seen to lead to a contradiction), and hence contains at least two linear triples. If the triple p_3, p_4, p_5, common to both quadruples, is linear, then the contradiction sought is obtained, for in this case both quadruples would be d-cyclic, and hence all five quadruples contained in the five points would be d-cyclic. But the circle has the congruence order four, and hence if all the quadruples formed from the five points are d-cyclic, the five points are themselves d-cyclic, and not pseudo d-cyclic, as supposed.

Let us select, then, another triple, say p_1, p_4, p_5 to be linear.* Then the

* If either of the remaining triples p_1, p_3, p_5 or p_1, p_3, p_4 is chosen linear, a contradiction due to the presence of diametral points is obtained in a similar manner.

quadruple p_1, p_3, p_4, p_5 is d-cyclic, and hence contains at least one other linear triple. Since we have already investigated the case in which the triple p_3, p_4, p_5 is linear, we suppose that one of the triples p_1, p_3, p_4; p_1, p_3, p_5 is linear. If the first of these triples is linear, then the quadruple p_1, p_2, p_3, p_4 has three of its triples linear, while if the second triple is linear, the quadruple p_1, p_2, p_3, p_5 has three of its triples linear. Now if either of the two remaining fourth triples, namely, p_2, p_3, p_4 and p_2, p_3, p_5, is linear, then the quadruple p_2, p_3, p_4, p_5 is d-cyclic and, as we have seen above, all the quadruples are d-cyclic, and the contradiction is obtained. We suppose, then, that neither of these two triples is linear, and we show that this is also impossible; for if a quadruple is d-cyclic and has three of its triples linear, while its fourth triple is not linear, then the quadruple must have two diametral points.* Hence either the quadruple p_1, p_2, p_3, p_4 or the quadruple p_1, p_2, p_3, p_5 has two diametral points, which is impossible, since by lemma 1 the quintuple, being proper, does not contain two diametral points.†

Thus, the assumption that a proper pseudo d-cyclic quintuple contains a linear triple is seen to lead to a contradiction, and the lemma is proved.

We are now in a position to prove the fundamental theorem for proper pseudo d-cyclic sets containing exactly five points. This theorem we state as follows:

Fundamental Theorem for proper pseudo d-cyclic quintuples. A proper pseudo d-cyclic quintuple is equilateral; that is, each two pairs of its points are congruent.

By lemma 2 we know that the set does not contain any linear triple.‡ Therefore, each of the five quadruples contained in the five points is a proper pseudo d-cyclic quadruple. Applying the theorem of section 3 to each of these five quadruples, we obtain immediately that all of the ten distances determined by the five points are equal, and an inspection of the relation satisfied by each of the ten d-cyclic triples shows that each distance is equal to $2d/3$. Hence, the theorem is proved.

Remark 1. We note, again, that the above theorem was proved for pseudo d-cyclic quintuples not containing a convex tripod nor a pseudo-linear quad-

* To prove this it is necessary to examine twenty-seven cases obtained by considering the possible ways in which the three triples can be linear.

† A similar method yields the desired contradiction in case the quadruple p_2, p_3, p_4, p_5 is chosen as containing a linear triple.

‡ Evidently, this is necessary for the validity of the theorem, for since the line does not contain an equilateral triangle, if one of the triples is linear, the three points forming the triple can not be equilateral.

ruple; that is, for proper pseudo d-cyclic quintuples. The following table furnishes an example of a pseudo d-cyclic quintuple that is *not* equilateral, but the set is not proper, since it contains convex tripods:

	p_1	p_2	p_3	p_4	p_5
p_1		$d/3$	$d/3$	$d/3$	$2d/3$
p_2	$d/3$		$2d/3$	$2d/3$	$d/3$
p_3	$d/3$	$2d/3$		$2d/3$	$d/3$
p_4	$d/3$	$2d/3$	$2d/3$		$d/3$
p_5	$2d/3$	$d/3$	$d/3$	$d/3$	

where the number appearing at the intersection of the i-th row and j-th column is the distance p_ip_j. It is easily verified that each three of the five points is d-cyclic, but the quadruple p_1, p_2, p_3, p_4 is seen to form a convex tripod, and is therefore not d-cyclic. Then, *a fortiori*, the five points are not d-cyclic, and are hence, pseudo d-cyclic. The triple p_2, p_3, p_4 is the only non-linear d-cyclic triple contained in the five points.

Remark 2. *A necessary and sufficient condition that a pseudo d-cyclic quintuple contain a linear triple is that the set contain a convex tripod, or a pseudo-linear quadruple.*

The sufficiency of the condition is evident, for if the set contains a convex tripod, then three of its triples are linear, and if it contains a pseudo-linear quadruple then four of its triples are linear. The necessity of the condition is seen to follow from the results just obtained, for if a pseudo d-cyclic quintuple contains a linear triple, the quintuple can not be proper (since by the theorem, if it were proper it would be equilateral, and hence would not contain a linear triple). Hence the set is improper; that is, it contains a convex tripod, or a pseudo-linear quadruple.

5. *Proper pseudo d-cyclic sets of n points.* We are now in a position to prove by induction the fundamental theorem characterizing proper pseudo d-cyclic sets containing more than four points. We state the theorem:

THEOREM. *A proper pseudo d-cyclic set containing more than four points is equilateral.*

We have proved the theorem for the case $n = 5$. Let us assume that the theorem is true for $n = k$, $k > 4$. We show that this implies the validity of the theorem for $n = k + 1$.

Consider, then, a proper pseudo d-cyclic set of $k + 1$ points. At least one of the sets of k points contained in these $k + 1$ points is pseudo d-cyclic;

for if every set of k points contained the $k+1$ points is d-cyclic, then, since $k>4$, every set of four points contained in the $k+1$ points is, *a fortiori*, d-cyclic. But since the circle has the congruence order four, it would follow that the set of $k+1$ points is d-cyclic, and not pseudo d-cyclic as assumed in the hypothesis. We may so label the $k+1$ points that the set

$$p_1, p_2, p_3, \cdots, p_{k-1}, p_k$$

is pseudo d-cyclic. Since the set of $k+1$ points is proper, the above pseudo d-cyclic set is also proper, and hence, by hypothesis, this set is equilateral, with each of the $(1/2)k(k-1)$ distances determined by these k points equal to $2d/3$.

We now show that the set of $k+1$ points must contain at least one other pseudo d-cyclic set of k points. To establish this, we suppose that

$$p_1, p_2, p_3, \cdots, p_{k-1}, p_k$$

is the only pseudo d-cyclic set contained in the $k+1$ points. Then the remaining k sets of k points into which the $k+1$ points of the set can be arranged are all d-cyclic, and hence all of the quadruples contained in these k sets are d-cyclic. Now these k sets may be obtained by omitting, in turn, from the $k+1$ points one of the points of the set $p_1, p_2, \cdots, p_{k-1}, p_k$. The k sets so obtained may be ordered by agreeing to call the i-th set that set in which the point p_i does not appear, $(i=1,\cdots,k)$. Then all of the quadruples contained in the pseudo d-cyclic set $p_1, p_2, \cdots, p_{k-1}, p_k$ that are independent of p_i are contained in the i-th one of these k sets. Thus every quadruple that can be formed from the k points $p_1, p_2, \cdots, p_{k-1}, p_k$ is to be found in one of the k d-cyclic sets, and hence is itself d-cyclic. But this is impossible, for if every quadruple contained in the set $p_1, p_2, \cdots, p_{k-1}, p_k$ is d-cycle, then the set is d-cyclic, and not pseudo d-cyclic as was supposed. Therefore, at least one other set of k points contained in the $k+1$ points of our set is pseudo d-cyclic.

To fix the ideas, let us suppose that the other pseudo d-cyclic set whose existence we have demonstrated above is the set $p_2, p_3, \cdots, p_k, p_{k+1}$. (If some other set is pseudo d-cyclic, the argument is, of course, entirely similar.) Then this set being proper is, by hypothesis, equilateral, and each of the distances determined by the k points is equal to $2d/3$. But we have already seen that each of the distances determined by the points $p_1, p_2, \cdots, p_{k-1}, p_k$ is equal to $2d/3$. Hence, of the $(1/2)k(k+1)$ distances determined by the $k+1$ points all are seen to be equal to $2d/3$ with the exception of the distance p_1p_{k+1}, which does not enter into these two sets.

To determine this distance, consider any triple containing p_1p_{k+1}, say the triple p_1, p_2, p_{k+1}. It is clear that this triple is not linear, for if it is linear, we have

$$(p_1p_2 + p_2p_{k+1} - p_1p_{k+1})(p_1p_2 - p_2p_{k+1} + p_1p_{k+1}) \times (p_1p_2 - p_2p_{k+1} - p_1p_{k+1}) = 0$$

and upon substituting for p_1p_2 and p_2p_{k+1} their equal, $2d/3$, we obtain $p_1p_{k+1} = 4d/3$, which is impossible. Hence the triple is not linear, and we have

$$p_1p_2 + p_2p_{k+1} + p_1p_{k+1} = 2d,$$

from which it is immediate that $p_1p_{k+1} = 2d/3$. Thus each distance has been shown equal to $2d/3$, and hence the set is equilateral.

Corollary 1. *A proper pseudo d-cyclic set of n points does not contain any linear triple.*

Corollary 2. *A necessary and sufficient condition that a pseudo d-cyclic set of n points* $(n \geq 4)$ *contains a linear triple is that the set be improper.*

Corollary 3. *A proper pseudo d-cyclic set of n points does not contain any diametral points.*

The Rice Institute.

ON SUBSETS OF A CONTINUOUS CURVE WHICH LIE ON AN ARC OF THE CONTINUOUS CURVE.*

By Edwin W. Miller

Introduction. In 1919 R. L. Moore and J. R. Kline gave the following *Theorem:* †

If M is a closed and bounded plane point set, in order that M be a subset of an arc (of the plane) it is necessary and sufficient that the following two conditions be satisfied:

1. *Every component of M is an arc or a point.*

2. *No interior point of an arc-component, t, of M is a limit point of $M - t$.*

As the plane is a particular example of a continuous curve,‡ the problem considered by Moore and Kline appears as a special case of the following general problem:

If S is a continuous curve in n-dimensional Euclidian space, E_n, and M is a closed and bounded subset of S, what conditions are necessary and sufficient that M be a subset of an arc of S?

It is with this problem that the present paper is chiefly concerned. The restriction that M be closed is not a vital one, as any set M is a subset of an arc if and only if the same is true of its closure $\bar{M}$.

Part I is given over to a study of the Moore-Kline conditions, which are of exceptional interest because of their purely internal character; that is to say, these conditions may be defined without reference to the imbedding space. As the proof of the above mentioned theorem of Moore and Kline makes use of properties of E_2 which are not properties of $E_n (n > 2)$ a different procedure from that pursued by Moore and Kline is adopted to obtain the result that the Moore-Kline theorem still holds if M is any bounded and closed subset of $E_n (n > 2)$.

In Part II a theorem is given which reduces the general problem for any continuous curve in E_n to the following problem: If a and b are two

* The chief results of Part I of this paper were presented under separate title to the American Mathematical Society on November 30, 1928. Part II was presented to the Society on April 30, 1930.

† *Annals of Mathematics*, Vol. 20 (1919), pp. 218-223.

‡ Throughout this paper the term "continuous curve" is used to mean a locally connected (connected im kleinen) continuum.

points of a cyclicly connected * continuous curve S, and M is a closed subset of S, under what conditions is M a subset of an arc in S whose end points are a and b? This problem remains unsolved as far as the present paper is concerned. However, with the aid of the above mentioned theorem complete solutions of the problem are obtained for certain important classes of continuous curves.

I.

THEOREM 1. *In $E_n (n > 2)$ in order that the closed and bounded set M be a subset of an arc it is necessary and sufficient that*

1. *Each component of M is either an arc or a point.*

2. *If t is an arc-component of M, then no interior point of t is a limit point of $M - t$.*

Proof. The conditions are evidently necessary.

They are also sufficient. L. W. Cohen has shown † that any set M in a separable metric space is homeomorphic with a subset of the linear continuum if and only if four conditions are satisfied, two of which conditions are the conditions given above, and the remaining two of which are:

3. If p is a point-component of M then $dim_p M = 0$; ‡ and

4. If a is an end point of an arc-component t of M then $dim_a (M - t + a) = 0$.

It will be shown that for a closed and bounded set M in E_n (the particular separable metric space here considered) condition (3) is satisfied and that (4) follows from (2). In fact, let F be a closed and bounded set in E_n, K a component of F, ϵ any positive number and H the set of all points of F which are at a distance $\geqq \epsilon$ from K. The set H, as is readily shown, is closed. It follows from a theorem of Miss Anna Mullikin § that F is the sum of two mutually exclusive closed sets F_1 and F_2 such that $F_1 \supset K$ and $F_2 \supset H$. If K is a single point we have the result that a closed set in E_n is 0-dimensional at any one of its point-components. Thus (3) is at once seen to be satisfied. From (2) it follows that if t is any arc-component of M and a and b are the end points of t, then $M - t + a + b$ is a closed set. The point a is evidently a point-component of $M - t + a + b$. Hence $dim_a (M - t$

* A continuous curve S is *cyclicly connected* if every two points of S lie on some simple closed curve which is a subset of S.

† *Fundamenta Mathematicae*, Vol. 14 (1929), pp. 281-303.

‡ For definitions of dimension, see Menger, *Dimensionstheorie* (Leipzig, 1928).

§ *Transactions of the American Mathematical Society*, Vol. 24 (1922), pp. 144-162.

$+a+b)=0$. But $dim_a(M-t+a+b)=dim_a(M-t+a)$. Thus (4) is also satisfied. The set M, then, is homeomorphic with some subset $\bar{M}$ of the linear continuum. Since M is closed and bounded, $\bar{M}$ is a subset of some linear interval I. Let $\bar{a}$ be the first point of $\bar{M}$ on I and $\bar{b}$ the last point of M on I, and consider the interval $(\bar{a},\bar{b})$. Let us arrange the open intervals complementary to $\bar{M}$ in a sequence $(\bar{a}_i,\bar{b}_i)$. It will be shown that corresponding to every i there is an arc N_i in E_n whose end points, a_i and b_i, are the points of M which correspond to $\bar{a}_i$ and $\bar{b}_i$, and which is such that $W=M+\sum_{i=1}^{\infty}N_i$ is an arc from a to b, the points of M corresponding to $\bar{a}$ and $\bar{b}$.

For this purpose we shall extend the following

Theorem of J. R. Kline: * *In E_n $(n\geqq 2)$ if H is the sum of a countable infinity of closed sets M_i, such that no $M_i=E_n$, and no M_i disconnects any domain, then if D is a domain, $D-D\cdot H$ is arc-wise connected.*

THEOREM A. *Under the conditions of the above theorem if a and b are any two points of D, there is an arc ab such that $ab-a-b\subset D-D\cdot H$.*

Proof. If $a\epsilon D\cdot H$, then a is a limit point of $D-D\cdot H$. In fact, since the closed set $M_i\neq E_n$, and disconnects no domain of E_n, it is easily shown to be nowhere dense. Hence H is of the first category of Baire, so that a is a limit point of $D-D\cdot H$. We can evidently find a sequence of distinct points $\{a_n\}$ such that $\lim a_n=a$, and such that the sphere

$$S_n=S(a_n,2d(a_n,a_{n+1}))$$

is a sub-domain of D. By the above theorem of Kline, $S_n-S_n\cdot H$ contains an arc a_na_{n+1}. The point set $a+\sum_{n=1}^{\infty}a_na_{n+1}$ is evidently a continuous curve containing a_1 and a. There is, then, an arc $a_1a\subset a+\sum_{n=1}^{\infty}a_na_{n+1}$ (R. L. Moore) †. Evidently, except for the point a, the arc $a_1a\subset D-D\cdot H$. From this result Theorem A follows readily.

Now, the set T consisting of all point-components of M together with the end points of all arc-components is clearly a closed, totally disconnected set. Furthermore, as Cohen shows,‡ the arc-components of M are countable.

* *Bulletin of the American Mathematical Society*, Vol. 23 (1917), pp. 290-292. The conditions here imposed on H are those given by R. L. Wilder, *ibid.*, Vol. 23 (1927), p. 388.

† See, for instance, R. L. Moore, *Bulletin of the American Mathematical Society*, Vol. 23 (1917), pp. 233-236.

‡ *Loc. cit.*

As any closed totally disconnected set is 0-dimensional and any arc is 1-dimensional,* M is the sum of a countable family of closed sets each of dimension $\leqq 1$. The set $M = T + \sum_{i=1}^{\infty} M_i$. Urysohn has shown † that a closed set of dimension $\leqq n-2$ disconnects no domain of E_n. Since we are here concerned with $E_n (n \geqq 3)$, neither T nor any one of the arc-components M_i disconnects any domain of E_n. Evidently neither T nor any one of the sets $M_i = E_n$. Hence the conditions of Theorem A are satisfied. Thus, if $S_1 = S(a_1, 2d(a_1, b_1))$ there is an arc N_1 from a_1 to b_1 such that

$$N_1 - a_1 - b_1 \subset S_1 - S_1 \cdot M.$$

The set $M + N_1$ satisfies the conditions of Theorem A. Thus, if

$$S_2 = S(a_2, 2d(a_2, b_2))$$

there is an arc N_2 from a_2 to b_2 such that

$$N_2 - a_2 - b_2 \subset S_2 - S_2 \cdot (M + N_1).$$

In general, the set $M + \sum_{i=1}^{n-1} N_i$ satisfies the conditions of Theorem A. Hence if $S_n = S(a_n, 2d(a_n, b_n))$ there is an arc N_n from a_n to b_n such that

$$N_n - a_n - b_n \subset S_n - S_n \cdot (M + \sum_{i=1}^{n-1} N_i).$$

It will now be shown that $W = M + \sum_{i=1}^{\infty} N_i$ is an arc.

As $\bar{M}$ is a closed bounded set the mapping of $\bar{M}$ on to M is uniformly continuous. Since the intervals $(\bar{a}_n, \bar{b}_n)$ do not overlap and all lie on a finite interval, $\lim_n d(\bar{a}_n, \bar{b}_n) = 0$. It follows that $\lim_n d(a_n, b_n) = 0$. Therefore $\lim_n \operatorname{diam}(S_n) = 0$ and as $N_n \subset S_n$, $\lim_n \operatorname{diam}(N_n) = 0$. Now let p be a limit point of W. If p is a limit point of M then $p \epsilon W$. If p is a limit point of $\sum_{i=1}^{\infty} N_i$, it either belongs to a particular N_i and therefore to W, or else it is a limit point of a set of points only a finite number of which belong to any given N_i. In this case, since $\lim_n \operatorname{diam}(N_n) = 0$, it is a limit point of the end points of the arcs N_i. But the end points of N_i are points of M. Thus W is closed.

If W is not connected it is the sum of two mutually exclusive non-vacuous closed sets Q_1 and Q_2. Not all the points of M belong to Q_1 (for example

* See Menger, *loc. cit.*

† *Fundamenta Mathematicae*, Vol. 8 (1926), p. 311.

for then we would have $Q_2 = 0$. Let us put, then, $M \cdot Q_1 = M'$ and $M \cdot Q_2 = M''$. The set $M = M' + M''$ where M' and M'' are both closed non-vacuous sets. This separation of M determines a separation of $\bar{M}$ into two mutually exclusive closed non-vacuous sets $\bar{M}'$ and $\bar{M}''$. Let $\bar{p}$ be any point of $\bar{M}'$. As $\bar{p}$ is not a limit point of $\bar{M}''$, and as there is at least one point of $\bar{M}''$ either to the left of $\bar{p}$ or to the right,—say to the right, there will be a first point $\bar{q}$ of the closed set $\bar{M}''$ to the right of $\bar{p}$. Since $\bar{q}$ is not a limit point of $\bar{M}'$ and since there is at least one point of $\bar{M}'$ to the left of $\bar{q}$, viz. $\bar{p}$—there is a last point $\bar{p}'$ of the closed set $\bar{M}'$ to the left of $\bar{q}$. Now $\bar{p}'$ and $\bar{q}$ are the end points of an interval complementary to $\bar{M}$. Therefore there is an arc N_i whose end points are p' and q. Hence $p' + q \subset Q_1$, or else $p' + q \subset Q_2$. It follows that $\bar{p}' + \bar{q} \subset \bar{M}'$ or else $\bar{p}' + \bar{q} \subset \bar{M}''$. This is in contradiction with the way in which $\bar{p}'$ and $\bar{q}$ were chosen.

We now show that $W - p$ is not connected if $W - (a + b) \supset p$.

Case 1. The point p belongs to M.

Let $W - p = W_1 + W_2$, where W_1 consists of all points x of M such that $\bar{x}$ is to the left of $\bar{p}$, together with all points x of W which are interior points of arcs N_i such that $\bar{a}_i$ and $\bar{b}_i$ are not to the right of $\bar{p}$; and where $W_2 = (W - p) - W_1$. The set $W_1 \neq 0$ since $W_1 \supset a$, and $W_2 \neq 0$ since $W_2 \supset b$. Now, let $q_1 \epsilon W_1$. There exists a sphere $S(q_1, d)$ such that $S(q_1, d) \cdot M \cdot W_2 = 0$ since the only limit point of the points of $\bar{M}$ to the right of $\bar{p}$ which is not a point of this set is the point $\bar{p}$. Since $\lim_i \operatorname{diam}(N_i) = 0$, there is a positive integer r such that if $i > r$, then $\operatorname{diam}(N_i) < d/2$. Hence if $i > r$, and if $N_i \subset W_2 + p$, then $S(q_1, d/2) \cdot N_i = 0$. Let F consist of all points z such that $z \epsilon N_i$ where $i \leqq r$. The set F is closed and $F \cdot q_1 = 0$. Since q_1 is neither a limit point of $M \cdot W_2$ nor of $W_2 \cdot \sum_{i=1}^{\infty} N_i$, it is not a limit point of W_2. Similarly no point of W_2 is a limit point of W_1.

Case 2. p is not a point of M.

Then p is a point of some N_i,—say N_k. If $a_k \neq a$, and $b_k \neq b$, then a_k and b_k determine separations: $W - a_k = W_1^{(a_k)} + W_2^{(a_k)}$, and $W - b_k = W_1^{(b_k)} + W_2^{(b_k)}$ in virtue of the result of Case 1. Denote by $a_k p$ the arc of N_k from a_k to p and by pb_k the arc of N_k from p to b_k, and put $W - p = K_1 + K_2$; where $K_1 = (W_1^{(a_k)} + a_k p - p)$ and $K_2 = (pb_k - p + W_2^{(b_k)})$. Since $K_1 \subset W_1^{(b_k)}$ it contains no limit point of $W_2^{(b_k)}$. It evidently contains no limit point of $pb_k - p$. Since $K_2 \subset W_2^{(a_k)}$ it contains no limit point of $W_1^{(a_k)}$. It evidently contains no limit point of $a_k p - p$. Thus $W - p$ is the sum of two mutually separated sets. In case $a_k = a$ or $b_k = b$, we can effect

a separation of $W - p$ in a similar fashion. This concludes the proof that W is an arc.

From Theorem 1 and the Moore-Kline theorem we have at once the following

COROLLARY: *If $M \subset E_n$ and $E_n \subset E_m$, where E_n and E_m are Euclidian spaces of n and m dimensions, then if M is a subset of an arc in E_m it is also a subset of an arc in E_n.*

THEOREM 2. *If M is a closed and bounded subset of a (connected) domain D of E_n, and if M satisfies the conditions of Moore and Kline, then there is an arc A such that $M \subset A \subset D$.*

Proof. By Theorem 1 there is an arc A' in E_n which contains M. If $q \epsilon A' \cdot D$, and C_q is the component of $A' \cdot D$ determined by q, only a finite number of distinct components C_q can contain points of M. For suppose there is an infinity of such components — $C_1, C_2, \cdots, C_n, \cdots$, and let $a_n \epsilon C_n \cdot M$. The points a_n are all distinct, and since M is bounded, the set $[a_n]$ has at least one limit point, p. Since M is closed, $p \epsilon M$ and therefore $p \epsilon D$. The component C_p is evidently a closed, half-closed or open arc, t.*

If t is a closed arc then evidently $A' \subset D$. If t is a half-closed arc denote $t +$ both its end points by t'. That end point a of t' which belongs to t is evidently an end point of the arc A'. In this case, then, p is either an end point of A' or an inner point of t'. But, clearly, then, since A' is an arc, p is not a limit point of $A' - t'$ and therefore not of $[a_n]$. Finally, if t is an open arc, and t' denotes $t +$ its end points, then p is an inner point of t'. But then, p is not a limit point of $A' - t'$, and therefore is not a limit point of $[a_n]$. Hence only a finite number of components C_q of $A' \cdot D$ contain points of M. Since M is a closed subset of D it is apparent that each C_q contains a closed sub-arc containing all the points of $M \cdot C_q$. Thus there exists a finite number of mutually exclusive arcs $A_1, \cdots A_m$ such that $M \subset \sum_{i=1}^{m} A_i \subset D$. It is an easy consequence of the Wilder accessibility theorem † that in E_2 the finite set of arcs $\sum_{i=1}^{m} A_i$ is contained in an arc of D. For E_n with $n > 2$ successive application of Theorem A leads to a like result.

THEOREM 3. *If D is a domain of E_n, and M is a closed bounded subset*

* An arc is said to be closed if it contains both end points; half-closed if it contains just one end point, and open if it contains neither end point.

† R. L. Wilder, *Fundamenta Mathematicae*, Vol. 7 (1923), pp. 340-377.

of D which satisfies the Moore-Kline conditions, and if p_1 and p_2 are distinct points of M, then there is an arc p_1p_2 such that $M \subset p_1p_2 \subset D$ if and only if.

(1) *p_1 and p_2 are point-components of M; or*

(2) *p_1 and p_2 are end points of distinct arc-components t_1 and t_2 of M such that p_1 is not a limit point of $M - t_1$ and p_2 is not a limit point of $M - t_2$; or*

(3) *one of the points p_1, p_2 is a point-component of M and the other is an end point of an arc-component t of M, and is not a limit point of $M - t$.*

Proof. Since no interior point of an arc-component t of M can be an end point of an arc which contains M, and since the same is true of an end point of an arc-component t of M which is a limit point of $M - t$, the conditions are at once seen to be necessary.

We now show that the conditions are also sufficient. By Theorem 2 there is an arc A such that $M \subset A \subset D$. We may suppose that the end points a_1 and a_2 of A are points of M. Let us first assume that neither a_1 nor a_2 is p_1 or p_2, and consider the case $n = 2$. On the basis of well-known results it is easy to show that there exists a simple closed curve J such that if I is the interior of J, then $J + I \subset D$ and $A \subset J$. If either p_1 or p_2 is an end point of an arc-component t of M and is not a limit point of $M - t$, or else is a point-component of M and is not a limit point of M from both sides on A, the simple closed curve J evidently contains an arc A' which contains M and has one of the two points, p_1 and p_2,—let us say, p_1, as an end point. If both of the points p_1 and p_2 are point-components of M and limit points of M from both sides on A, we shall construct an arc A' which contains M and has p_1 as one of its end points as follows: We may evidently suppose that the order of the points a_1, p_1, p_2, and a_2 is a_1, p_1, p_2, a_2. Let us choose two sequences of sub-arcs of A complementary to M, with end points x_n and y_n and x'_n and y'_n respectively, so that

$$a_1 < x'_{n-1} < y'_{n-1} < x'_n < y'_n < p_1 < x_n < y_n < x_{n-1} < y_{n-1} < p_2,$$

and such that $\lim x'_n = \lim y'_n = \lim x_n = \lim y_n = p_1$.

There is an arc α_1, from x'_1 to x_1 such that $\alpha_1 - x'_1 - x_1 \subset I$. The arc α_1 forms with the sub-arc x'_1x_1 of A a simple closed curve J_1 whose interior $I_1 \subset I$. There is an arc α'_1 from y_2 to y'_1 such that $\alpha'_1 - y_2 - y'_1 \subset I_1$. The arc α'_1 forms with the sub-arc $y_2y'_1$ of A a simple closed curve J'_1 whose interior $I'_1 \subset I_1$.

We proceed in this way to construct arcs α_2 from x'_2 to x_2, α'_2 from y_3 to y'_2, α_3 from x'_3 to x_3, and so on. Now, it is possible to choose the arcs α_n

and α'_n so that $\lim \operatorname{diam}(\alpha_n) = \lim \operatorname{diam}(\alpha'_n) = 0$. In fact, let us prove the following

LEMMA. *If x and y are two points of a simple closed curve J, whose interior is I, and xy is either arc of J from x to y, then there exists an arc α from x to y such that $\alpha - x - y \subset I$, and $\operatorname{diam}(\alpha) < \operatorname{diam}(xy) + \epsilon$, where ϵ is a preassigned positive number.*

Proof. Let D denote the set of points of I at a distance $< \epsilon/2$ from the arc xy. It is easily shown that D is an open and connected subset of I. The points x and y lie on the boundary of D and neither belongs to any continuum of condensation of the boundary. Hence, by the Wilder accessibility theorem, x and y are both accessible from D. It follows at once that there is an arc α in I from x to y such that $\operatorname{diam}(\alpha) < \operatorname{diam}(xy) + \epsilon$.

We will suppose, then, that the arcs α_n and α'_n are chosen so that $\lim \operatorname{diam}(\alpha_n) = \lim \operatorname{diam}(\alpha'_n) = 0$. Then the point set A', consisting of the arc $y_1p_2a_2$ of A, that arc x'_1a_2 of J which does not contain p_1, the arcs α_n and α'_n, and the sub-arcs x_ny_{n+1} and $y'_nx'_{n+1}$ of A $(n = 1, 2, \cdots)$ is, as can easily be shown, an arc from p_1 to y_1 which contains M and lies in D.

An argument which is only a slight modification of the one given above suffices to show that starting with the arc A' we can construct an arc A'' of D which contains M and has p_1 *and* p_2 as its end points. This completes the proof for the case $n = 2$.

For the case $n > 2$, our procedure is similar. By Theorem A, the end points of the arc A are accessible from D. There is then a simple closed curve J such that $A \subset J \subset D$. As in the case $n = 2$, we are reduced to the supposition that p_1 is a point-component of M which is a limit point of M from both sides on A. We choose two sequences of sub-arcs of A complementary to M as before, and construct arcs α_n and α'_n, all mutually exclusive, such that $\lim \operatorname{diam}(\alpha_n) = \lim \operatorname{diam}(\alpha'_n) = 0$. That there exist two such sequences of arcs α_n and α'_n is readily seen to be a consequence of Theorem A. The rest of the proof runs parallel to the proof for the case $n = 2$.

It is natural, in this connection, to consider the problem of determining conditions under which a closed point set M in E_n is a subset of a ray or an open curve.*

THEOREM 4. *A closed set M in E_n is a subset of a ray if and only if the two following conditions are satisfied:*

1. *Either (a) the components of M are all arcs and points, or (b) the*

* For definitions of these, see R. L. Moore, *Transactions of the American Mathematical Society*, Vol. 21 (1920), p. 347.

components of M are a single ray and a bounded collection of arcs and points; and

(2) *No interior point of an arc- or ray-component, t, is a limit point of $M - t$.*

The conditions are evidently necessary. Their sufficiency is proved as follows: Suppose, first, that 1(a) and 2 hold. Let $p \epsilon E_n - M$, and let $S = S(p, d)$ be such that $S \cdot M = 0$. If S is taken as the sphere of inversion, the set M maps into a set $M' \subset S$. By familiar properties of the transformation of inversion, the set $M' + p$ is closed and satisfies the Moore-Kline conditions. The point p is a point-component of $M' + p$. Thus, by Theorem 3 there is an arc which contains $M' + p$, lies in S, and has p as one of its end points. The transform of this arc is a ray which contains M.

If 1(b) and 2 hold, $M' + p$ is again a set satisfying the Moore-Kline conditions. In this case p is an end point of an arc-component t of $M' + p$ and is not a limit point of $(M' + p) - t$. Again, Theorem 3 gives us the existence of an arc in S which contains $M' + p$ and has p as one end point. The transform of this arc is a ray which contains M.

It is here to be noted that G. T. Whyburn has given conditions which are necessary and sufficient that a closed set in E_2 be a subset of an open curve.* Whyburn's proof depends upon the Moore-Kline theorem. Using Theorem 1 of this paper in place of the Moore-Kline theorem, the result can be extended to $E_n (n > 2)$.

II.

In the first part of this section it will be shown that the Moore-Kline conditions constitute a solution of the problem of this paper only for a rather limited class of continuous curves.

THEOREM 5. *If S is a plane closed set, and if every one of its bounded closed subsets which satisfies the Moore-Kline conditions is a subset of an arc in S, then S is either the whole plane or a simple continuous curve.*

Proof. To abbreviate we will use the expression "a set E" to mean any bounded closed subset of S which satisfies the Moore-Kline conditions.

Since any pair of points of S is a set E, S is arc-wise connected and therefore connected. Since S is closed, it is a continuum. S is, in fact, a continuous curve. For if S is not a continuous curve, then in virtue of a

* *Transactions of the American Mathematical Society*, Vol. 29 (1927), pp. 746-754; Theorem 5.

theorem of R. L. Wilder's * there is a point p of S and a circular neighborhood K of p such that p is a sequential limit point of a set of points $[p_n]$ of S each of which lies in a different quasi-component of $S \cdot K$. Evidently, $p + \sum p_n$ is a set E, and if t were an arc of S containing it, there would be in $S \cdot K$ a sub-arc t' of t such that $t' \supset p + \sum_{n=k}^{\infty} p_n (k \geqq 1)$. This, of course, contradicts the fact that no two points p_k and p_j lie in the same component of $S \cdot K$, and consequently S is a continuous curve.

It will now be shown that S is either the whole plane or else a simple continuous curve. Let us suppose, first, that S contains no simple closed curve. If S contained a point p which separated S into three mutually separated sets, S_1, S_2, and S_3, then if $p_i \epsilon S_i$ $(i = 1, 2, 3)$, the set $\sum_{i=1}^{3} p_i$ would be a set E which is clearly not a subset of any arc in S. Hence, no point p of S separates S into more than two components.

If S is bounded, it has at least two non-cut points, q_1 and q_2.† Let t be an arc of S from q_1 to q_2. If $t \neq S$, there is a component C of $S - t$ which has—since S is an acyclic continuous curve—a unique limit point q_3 in t. The set $S - q_3$, as is easily shown, would consist of three mutually separated sets. We have seen that this is impossible. Hence, if S is bounded, it is an arc.

Let us now suppose that S is unbounded and that it contains a non-cut point, p. The point p is the end point of a ray r of S.‡ If $r \neq S$, the unique limit point in r of a component C of $S - r$ would break up S into three mutually separated sets. Hence if S is unbounded and contains a non-cut point, it is a ray.

Let us suppose, finally, that S is unbounded and is disconnected by each one of its points. Thus if $p \epsilon S$, the set $S - p$ consists of two mutually separated connected sets, S_1 and S_2. The point p is a non-cut point of $S_i + p$ $(i = 1, 2)$, and by the argument just employed $S_i + p$ is a ray r_i, so that $S = r_1 + r_2$ is an open curve.

Now, let us suppose that S does contain a simple closed curve, J. Then S is cyclicly connected. For, suppose not. Then J determines a maximal

---

* *Proceedings of the National Academy of Sciences*, Vol. 15 (1929), pp. 614-621; Theorem 1.

† S. Mazurkiewicz, *Fundamenta Mathematicae*, Vol. 2 (1921), pp. 119-130.

‡ C. Kuratowski, *Fundamenta Mathematicae*, Vol. 3 (1922), pp. 59-64.

cyclic curve* N of S such that $N \neq S$. If C is any component of $S - N$ then C has just one limit point, x, in N.* If a and b are two points of N, then, since N is cyclicly connected, there is an arc t in N from a to b one of whose interior points is x.† If $y \epsilon C$, the set $t + y$ is a set E, but it not a subset of any arc in S. Hence S is cyclicly connected.

Since S is a plane cyclicly connected continuous curve, the boundary of any domain complementary to S is either a simple closed curve or an open curve.† If $S \neq E_2$, there is at least one domain complementary to S.

Case 1. There is a domain D complementary to S whose boundary is a simple closed curve, J, of S.

It is easily shown that D is either the exterior or the interior of J. Let us suppose that D is the interior of J. (The argument is essentially the same if D is the exterior of J.) If $J \neq S$, there is a component C of $S - J$ which lies in the exterior of J. Since S is cyclicly connected, C has at least two limit points in J. Using the Wilder accessibility theorem, it follows that there exists an arc α whose end points a_1 and a_2 lie on J and which is such that $\alpha - a_1 - a_2 \subset C$. Let a_1a_2 be that one of the two arcs of J from a_1 to a_2 which forms with α a simple closed curve whose interior lies exterior to J. Now let $r \epsilon J - a_1a_2$ and let a'_1 and a'_2 be points of a_1a_2 such that $a_1 < a'_1 < a'_2 < a_2$. Then the E-set $\alpha + a_1a'_1 + a_2a'_2 + r$ is clearly a subset of no arc in S. Thus $J = S$.

Case 2. There is no domain complementary to S whose boundary is a simple closed curve. We shall show that, under our supposition that S contains a simple closed curve, this case is impossible.

Let D be any domain complementary to S. The boundary, B, of D is an open curve. Since S contains a simple closed curve, $B \neq S$. There is, then, a component C of $S - B$. In the same way as in Case 1, we prove the existence of an arc α whose end points, a_1 and a_2, are points of B, and which is such that $\alpha - a_1 - a_2 \subset S - B$. Let $r \epsilon B - a_1a_2$ and let a'_1 and a'_2 be points of a_1a_2 such that $a_1 < a'_1 < a'_2 < a_2$. As D must be identical with that domain of the plane complementary to B which does not contain $\alpha - a_1 - a_2$, it is clear that the E-set $\alpha + a_1a'_1 + a'_2a_2 + r$ is a subset of no arc in S. Thus, Case 2 has been shown to be impossible.

---

* G. T. Whyburn, *Proceedings of the National Academy of Sciences*, Vol. 13 (1927), pp. 31-38.

† W. L. Ayres, *Bulletin de l'Académie Polonaise des Sciences et des Lettres* (1928); Theorem 3.

THEOREM 6. *If S is a closed set in a space E_n, and is of dimension 1 in at least one of its points, and if every bounded closed subset of S which satisfies the Moore-Kline conditions is a subset of an arc in S, then S is a simple continuous curve.*

Proof. As in the proof of the previous theorem, let us use the expression "a set E" to mean any bounded closed subset of S which satisfies the Moore-Kline conditions. As established in the previous theorem, S is a continuous curve, and if S contains no simple closed curve it is either an arc, a ray, or an open curve; and if S does contain a simple closed curve it is cyclicly connected.

We suppose, then, that S is cyclicly connected, and denote by p a point in which S is 1-dimensional. The point p lies on some simple closed curve J of S. Now if $J \neq S$, there is a component C of $S - J$. The component C has at least two limit points in J, since S is cyclicly connected. It follows that there is an arc in S which is a subset of $S - J$ except for one of its end points q which is a point of J distinct from p. Let $\sum$ be a sphere whose center is p and which is so small that q does not lie in $\sum$ or on its boundary. Since S is 1-dimensional in p there is a vicinity V of p such that $V \subset \sum$ and such that B, the boundary of V in S, is totally disconnected. As $B + p$ is a closed totally disconnected set and therefore a set E, it is a subset of an arc A of S. Now A cannot contain all the points of S outside V, for the point q lies outside V and is the junction point of three arcs which lie outside V. Therefore, there is a point h outside V which is not a point of A. If C_h is the component of $S - A$ which h determines, then C_h cannot have p as a limit point since p lies inside V, and h outside, and A contains the boundary B of V in S.

Now, there must be an arc in C_h from h to one of the end points a_1 or a_2, say, a_1—of A, for otherwise the E-set $A + h$ is a subset of no arc in S. Since S is cyclicly connected, C_h has another limit point in A. Let r be the last limit point of C_h on A in the order from a_1 to a_2. The point r is arc-wise accessible from C_h. For let a_1r be the arc of A from a_1 to r, and let $\{r_n\}$ be a sequence of points in C_h such that $\lim r_n = r$. The set $a_1r + [r_n]$ is an E-set and therefore a subset of an arc in S. Clearly, this arc contains an arc from a point r_n to r which lies in C_h except for r. Thus r is accessible from C_h. Let t_1 and t_r be mutually exclusive arcs of S, where a_1 is an end point of t_1, and r is an end point of t_r, and where $(t_1 - a_1) + (t_r - r) \subset C_h$. The point r must be a_2, for suppose not. Then the point set $t_1 + a_1r + t_r + a_2$ is a set E, but is clearly not a subset of any arc in S. Hence $r = a_2$.

If there were a component C_w of $S - A$ distinct from C_k, the point set $t_1 + a_1r + t_r + w$ would be an E-set which is clearly not a subset of any arc in S. Thus C_k is the only component of $S - A$, so that $S = A + C_k$. Since a_1 and a_2 are both accessible from C_k, there is an arc L from a_1 to a_2 such that $L - a_1 - a_2 \subset C_k$. Now, p is an interior point of A and is not a limit point of C_k. Therefore there is a whole arc x_1px_2 of A, no point of which is a limit point of C_k, and where the order on A of the points a_1, x_1, p, x_2 and a_2 is a_1, x_1, p, x_2, a_2. Let a_1x_1 and x_2a_2 be the arcs of A from a_1 to x_1 and from x_2 to a_2, respectively. Suppose now that $C_k - C_k \cdot L \neq 0$ and let $s \epsilon C_k - C_k \cdot L$. The point set $a_1x_1 + L + a_2x_2 + s$ is a set E, but is a subset of no arc in S. Hence $C_k = C_k \cdot L$ so that S is the simple closed curve $L + A$. We have shown, then, that if S contains a simple closed curve, S is a simple closed curve. This completes the proof of the theorem.

It is not apparent that in three dimensions the closed sets for which the Moore-Kline theorem is true fall into any easily classified category. For instance, besides the topological plane and sphere, and the simple continuous curves, the theorem is true in three dimensions for such sets as the cube together with its interior, any closed two-dimensional manifold, and various continua made up of combinations of closed manifolds.

We shall now consider the general problem of determining when a given closed subset M of a continuous curve S lies on an arc of S.

THEOREM 7. *If S is a bounded continuous curve, and M a closed subset of S which is not a subset of any one maximal cyclic curve of S, then M is a subset of an arc in S if and only if the following three conditions are satisfied:*

(1) *If C is a cyclic element* of S, at most two components of $S - C$ contain a point of M.*

(2) *If C is a maximal cyclic curve of S, and K_1 and K_2 are distinct components of $S - C$ such that $K_i \cdot M \neq 0$ $(i = 1, 2)$, and such that a_1, the unique limit point of K_1 in C is distinct from a_2, the unique limit point of K_2 in C, then there is an arc t in C from a_1 to a_2 such that $M \cdot C \subset t$.*

(3) *If C is a maximal cyclic curve of S, and K is the only component of $S - C$ such that $K \cdot M \neq 0$, and a is the unique limit point of K in C, then there exists an arc t in C with a as one end point and such that $C \cdot M \subset t$.*

Proof. The conditions are sufficient.

* See G. T. Whyburn, *American Journal of Mathematics*, Vol. 50 (1928), pp. 167-194.

Let W be the arc-curve $S(M)$.* The set W is a continuous curve* consisting of a certain collection of cyclic elements of S, and the cyclic elements of W are cyclic elements of S.† It will first be shown that if C is a cyclic element of W, then $W - C$ contains at most two components. Let C be any cyclic element of W. If K is a component of $W - C$, then $K \cdot M \neq 0$. For suppose the contrary and let a be the unique limit point of K in C, and let $x \epsilon K$. Since $K \subset W$, there are two points m_1 and m_2 of M which are the end points of an arc t in W which has x as an interior point. Not both of the sub-arcs m_1x and xm_2 of t can contain the point a. But m_1 and m_2 are both points of $W - K$ and a separates K from $W - K - a$. From this contradiction it follows that $K \cdot M \neq 0$.

Suppose, now, that $W - C$ contains three components K_i $(i = 1, 2, 3)$. Each component K_i is a subset of a component K'_i of $S - C$. In virtue of (1), two of these components are identical. We may suppose $K'_1 = K'_2$. Let $x \epsilon K_1 \cdot M$ and $y \epsilon K_2 \cdot M$. Now, $x + y \subset K'_1$, and there is an arc t in K'_1 from x to y.‡ Since $t \subset W$ and $K_1 + t + K_2$ is a connected subset of $W - C$, K_1 and K_2 lie in one component of $W - C$. As this contradicts the definition of K_1 and K_2, it follows that $W - C$ contains at most two components.

G. T. Whyburn has shown § that any bounded continuous curve, H, which is not cyclicly connected contains at least two nodes, a node being either an end point of H or a maximal cyclic curve N of H such that $H - I(N)$ is connected, $I(N)$ being the set of all points of N which are not limit points of $H - N$.

Since, by our hypothesis, M is not a subset of any one maximal cyclic curve of S, and since the cyclic elements of W are all cyclic elements of S, W cannot be cyclicly connected. Accordingly let N be a node of W. Suppose $W - N$ is not connected. Then N is evidently not a point since no end point of a continuous curve cuts the continuous curve. We have $W - N = W_1 + W_2$, where W_1 and W_2 are mutually separated. As we have already shown, W_1 and W_2 are connected. The set W_1 has a unique limit point a_1 and W_2 a unique limit point a_2 in N. Now $a_1 \neq a_2$, for if $a_1 = a_2$, the set $W - a_1$ contains three components $N - a_1$, W_1 and W_2. Evidently $W - I(N)$

---

* See W. L. Ayres, *Transactions of the American Mathematical Society*, Vol. 30 (1928), pp. 567-578.

† W. L. Ayres, *Transactions of the American Mathematical Society*, Vol. 31 (1929), pp. 595-612.

‡ See, for instance, R. L. Moore, *Bulletin of the American Mathematical Society*, Vol. 23 (1917), pp. 233-236.

§ See G. T. Whyburn, *American Journal of Mathematics*, *loc. cit.*

$= (W_1 + a_1) + (W_2 + a_2)$. Since W_1 and W_2 are mutually separated and since $a_1 \cdot \bar{W}_2 = a_2 \cdot \bar{W}_1 = 0$ it follows that $(W_1 + a_1)$ and $(W_2 + a_2)$ are mutually separated. As this contradicts the fact that N is a node, every node of W must be a non-cut element of W, so that W contains at least two non-cut elements N_1 and N_2.

Let X be any simple cyclic chain* in W from N_1 to N_2. It will be shown that $X = W$. Suppose, in fact, $X \neq W$, and let $p \epsilon W - X$. The point p determines a component C_p of $W - X$ which has a unique limit point q in X. Since N_1 and N_2 are non-cut elements of W, they are, in virtue of a theorem of Whyburn's,† end elements of W. In other words, q is not a point of either N_1 or N_2. Hence, q is a point of some cyclic element Q of X interior to X. Clearly, $W - Q$ contains three components determined by $N_1 - N_1 \cdot Q$, $N_2 - N_2 \cdot Q$, and C_p. As we have shown that this is impossible, it follows that W is the simple cyclic chain X.

Now, let t be an arc in W from a point of N_1 to a point of N_2. In virtue of a theorem of Whyburn's,‡ the arc t contains a point of every cyclic element of W, and in fact, if C is a maximal cyclic curve of W interior to X, then $t \cdot C$ is a sub-arc of t whose end points are the unique limit points in C of the two distinct components of $W - C$ and therefore of $S - C$. Let us now replace all such sub-arcs of t by the corresponding arcs whose existence is given by (2), and let us replace $t \cdot N_1$ and $t \cdot N_2$ by sub-arcs of N_1 and N_2 of the sort whose existence is given by (3). Since for every $\epsilon > 0$ there are only a finite number of maximal cyclic curves of W of diameter $> \epsilon$ it is readily shown that the resulting point set is an arc, A. As $M \subset A \subset W \subset S$, the conditions have been shown to be sufficient.

The conditions are also necessary.

Let t be an arc of S such that $t \supset M$, and let us suppose that there is a cyclic element C of S such that three components, S_1, S_2, and S_3 of $S - C$ each contain a point of M. Let $p_i \epsilon M \cdot S_i$, and let us suppose that the order of the points p_1, p_2, and p_3 on t is p_1, p_2, p_3. Since these three points lie in different components of $S - C$, the sub-arcs p_1p_2 and p_2p_3 of t each contain a point of C. Evidently, then, S_2 has at least two limit points in C. As this is impossible, the necessity of (1) is proved.

To prove the necessity of (2) let C be any maximal cyclic curve of S and suppose there exist components K_1 and K_2 of $S - C$ and points a_1 and a_2

* See G. T. Whyburn, *loc. cit.*

† *Loc. cit.*

‡ *Loc. cit.*

as specified in (2). Since each point a_1 and a_2 separates K_1 from K_2, t must contain a_1 and a_2. The sub-arc a_1a_2 of t is a subset of C, for otherwise there is determined a component of $S - C$ which has at least two limit points in C. If $(t - a_1a_2) \cdot C \neq 0$ either K_1 or K_2 would have two limit points in C. Hence $a_1a_2 = t \cdot C$. Since $t \supset M \supset C \cdot M$, the arc a_1a_2 of C must contain $C \cdot M$. The necessity of (3) follows in a similar way. This completes the proof of the theorem.

We shall now make certain applications of the theorem just proved.

THEOREM 8. *If the plane continuous curve S is the boundary of a (connected) domain D and M is any closed subset of S other than a simple closed curve, then M is a subset of an arc in S if and only if the following three conditions are satisfied:*

(1) *If F is a point or a simple closed curve of S, at most two components of $S - F$ contain a point of M.*

(2) *If J is a simple closed curve in S, and if there are two components of $S - J$ which contain a point of M and have distinct limit points a_1 and a_2 in J, then only one of the arcs of J from a_1 to a_2 contains a point of M.*

(3) *If J is a simple closed curve in S, and if just one component of $S - J$ contains a point of M, then the (unique) limit point in J of this component is not a limit point of $J \cdot M$ from both directions along J.**

Proof. The theorem is almost an immediate consequence of Theorem 7 and the fact that every maximal cyclic curve of a continuous curve which is the boundary of a plane domain is a simple closed curve.† It is clear that the above theorem is true for any continuous curve in E_n whose maximal cyclic curves are all simple closed curves.

THEOREM 9. *If S is an acyclic continuous curve in E_n and M is a closed subset of S, then in order that M be a subset of an arc in S it is necessary and sufficient that if p is a point of S, at most two components of $S - p$ contain a point of M.*

Proof. This theorem is an immediate consequence of Theorem 7 and the fact that every cyclic element of an acyclic continuous curve is a point.

Theorem 9 can evidently be made more general by letting S be *any continuous curve in E_n, M any closed subset of S which has no point on any simple closed curve of S, and p any cyclic element of S.*

---

* The phrase "from both directions along J" is immediately translatable into topological language.

† R. L. Wilder, *Fundamenta Mathematicae*, Vol. 7 (1923), pp. 340-377.

With a view to making another application of Theorem 7, the following theorem will be proved:

THEOREM 10. *If J is a simple closed curve in E_2, and I is its interior, and if F is a closed totally disconnected subset of $J + I$, then there is an arc in $J + I$ which contains F.*

Proof. Let us denote $F \cdot J$ by T. There is an arc AB of J which has T in its interior. Let C and D be two other points of J and let the order of the points A, B, C, and D on J be A, B, C, D, A, where T is contained in the arc ABC of J. Let a system of rectangular coördinates be given, and let A', B', C' and D' be the points $(0,0)$, $(1,0)$, $(1,1)$, and $(0,1)$, respectively. Let J' be the simple closed curve formed by the square $A'B'C'D'A'$, and let I' be its interior. There will be defined a $1 - 1$ bicontinuous mapping of $J + I$ into $J' + I'$ which has the properties:

(1) J corresponds to J'; the points of AB corresponding to the points of $A'B'$, A corresponding to A' and B to B'.

(2) The points of T correspond to irrational points of $A'B'$.

(3) The points of $F - T$ correspond to points of I' both of whose coördinates are irrational.

The proof that such a mapping exists is based on a lemma in a paper of R. L. Wilder's,* and for the meaning of the notation used in the next paragraph, the reader is referred to the paper in question.

There is a $1 - 1$ bicontinuous correspondence between the points of AB and the points of $A'B'$ in which A corresponds to A' and B to B', the rational points within $A'B'$ corresponding to points of AB which are end points of the arcs a_1, and the irrational points of $A'B'$ corresponding to points of AB which are end points of the arcs b_1. There is a like $1 - 1$ bicontinuous correspondence between the points of $A'D'$ and AD. These two correspondences induce like correspondences between the points of BC and $B'C'$, CD and $C'D'$. If $p \epsilon I$, there is just one arc t_1 of α_1, and just one arc t_2 of α_2 which contains p. That end point of t_1 which belongs to AB corresponds to a point x of $A'B'$, and that end point of t_2 which belongs to AD corresponds to a point y of $A'D'$. To the point p we let correspond the point (x, y) of I'. It is readily shown that the correspondence in question is $1 - 1$ bicontinuous and has the properties (1), (2), and (3) specified above.

Consider, now, the square J', and let F' and T' be the sets corresponding

---

* *Transactions of the American Mathematical Society*, Vol. 31 (1929), pp. 345-359; Lemma 6, p. 356.

to F and T. The set T' is a closed non-dense subset of $A'B'$. Let R_1 be the rectangle whose vertices are the points $(0, \frac{1}{2})$, $(1, \frac{1}{2})$, $(1, 1)$ and $(0, 1)$. The projection, p_1, on $A'B'$ of the mid-point of the base of R_1 is the point $(\frac{1}{2}, 0)$. Since p_1 is a rational point, it belongs to some sub-interval of $A'B'$ complementary to T'. Let $(a_1, 0)$ and $(b_1, 0)$ be any two rational points of this interval such that $a_1 <$ absc. of $p_1 < b_1$. Let J_1 be the simple closed curve obtained by deleting from R_1 the sub-interval of R_1 between $(a_1, \frac{1}{2})$ and $(b_1, \frac{1}{2})$ and adding the lines representing the ordinates of these two points and the interval, i_1, of $A'B'$ from $(a_1, 0)$ to $(b_1, 0)$. Denote, now, by R_2 the rectangle whose vertices are the points $(0, \frac{1}{2^2})$, $(a_1, \frac{1}{2^2})$, $(a_1, \frac{1}{2})$ and $(0, \frac{1}{2})$. The projection, p_2, on $A'B'$ of the mid-point of the base of R_2 is a rational point and therefore belongs to some sub-interval of $A'B'$ complementary to T'. Let $(a_2, 0)$ and $(b_2, 0)$ be any two rational points of this interval such that $a_2 <$ absc. of $p_3 < b_2$, and let us denote the interval of $A'B'$ from $(a_2, 0)$ to $(b_2, 0)$ by i_2. These points are related to R_2 in the same way that $(a_1, 0)$ and $(b_1, 0)$ are related to R_1. Let us, then, construct J_2 in relation to them and to R_2 in the same way that J_1 was constructed in relation to $(a_1, 0)$ and $(b_1, 0)$ and R_1. We now consider the rectangle R_3 whose vertices are the points $(b_1, \frac{1}{2^2})$, $(1, \frac{1}{2^2})$, $(1, \frac{1}{2})$ and $(b_1, \frac{1}{2})$. We define in a similar way a simple closed curve J_3. At the next stage of the process we have 2^2 rectangles to consider, then 2^3 rectangles, and so on.

Proceeding indefinitely in the manner indicated, it is clear that we obtain a sequence $\{J_n\}$ of simple closed curves such that:

(1) $J_n \cdot F' = 0 \qquad (n = 1, 2, \cdots)$.

(2) If H_n is the interior of J_n, then $H_i \cdot H_j = 0$ $(i \neq j)$ $(i, j = 1, 2, \cdots)$.

(3) $F' - T' \subset \sum_{n=1}^{\infty} H_n$.

(4) $\lim \operatorname{diam}(J_n) = 0$.

The set F'_n is a closed totally disconnected set. There is, then, in virtue of Theorem 2 of this paper an arc t_n in H_n such that $t_n \supset F'_n$. It is easy to show that the end points of this arc can be joined by arcs to a_n and b_n so as to give an arc t'_n such that $t'_n \supset F'_n$ and $t'_n - a_n - b_n \subset H_n$. It follows easily that the point set $\sum_{n=1}^{\infty} t'_n + (A'B' - \sum_{n=1}^{\infty} i_n)$ is an arc, t', whose end points are A' and B'. Evidently $F' \subset t' \subset J' + I'$. In virtue of the homeomorphism between $J + I$ and $J' + I'$, the theorem is established.

THEOREM 11. *Under the conditions of Theorem* 10, *if a and b are two*

points of F, then there is an arc in $J + I$ which contains F and has a and b as its end points.

The proof is based on Theorem 10 and uses the homeomorphism between $J + I$ and $J' + I'$ set up in the proof of that theorem.

It is possible to show that the arc of Theorem 11 can be so constructed that the only points it has in common with J are the points of F in J.

We shall now use Theorem 11 to make another application of Theorem 7.

THEOREM 12. *If S is a bounded plane continuous curve which does not cut the plane, and F is a closed totally disconnected subset of S, then in order that F lie on an arc of S, it is necessary and sufficient that for every cyclic element C of S at most two components of $S - C$ contain a point of F.*

Proof. Every maximal cyclic curve of S is a simple closed curve plus its interior.* If C is such that there are two points a_1 and a_2 of the sort specified under condition (2), Theorem 7, then in virtue of Theorem 11, there is an arc in C which contains the closed totally disconnected set $C \cdot F + a_1 + a_2$ and has a_1 and a_2 as its end points. Similarly, if condition (3), Theorem 7, applies, there is an arc in C which contains $C \cdot F + a$, and has a as one of its end points. Thus the conditions of Theorem 7 are all satisfied and Theorem 12 follows.

As a special case of the general problem of this paper it is of interest to consider the problem of determining necessary and sufficient conditions that a given finite subset of a continuous curve S lie on an arc of S. A condition which is sufficient but not necessary has been given by W. L. Ayres.† Evidently, if S is any set of points whatsoever and M is a subset of S consisting of just n points, then for M to lie on an arc of S it is *necessary* that if $p_1, \cdots p_i$ are any i points of S $(i < n - 1)$ at most $i + 1$ components of $S - (p_1 + \cdots + p_i)$ contain a point of M. This condition is clearly not sufficient for continuous curves in general.

We have found conditions which are at the same time necessary and sufficient that a given set of n points of a continuous curve S lie on an arc of S for the cases $n = 3$ and $n = 4$. These are embodied in the following two theorems whose proofs we do not feel it will be necessary to include.

THEOREM 13. *If S is a continuous curve in E_n and M is a subset of S consisting of just three points, then M is a subset of an arc in S if and only*

---

* See G. T. Whyburn, *American Journal of Mathematics*, *loc. cit.*

† This paper is not yet published. For abstract, see *Bulletin of the American Mathematical Society*, Vol. 35 (1929), p. 772.

if M contains a point which is not separated from both of the other points of M by any one point of S.

THEOREM 14.* *If S is a cyclicly connected continuous curve in E_n and M consists of four points of S, then M is a subset of an arc in S if and only if for every pair of points p_1 and p_2 of S at most three components of $S-(p_1+p_2)$ contain a point of M.*

Simple examples show that none of the evident generalizations of the conditions in Theorems 13 and 14 apply to the general case of *any* finite subset of a continuous curve.

* This theorem is a generalization of a theorem of W. L. Ayres, ***Bulletin de l'Académie Polonaise des Sciences et des Lettres*** (1928), Theorem 5.

NOTE ON THE CONDITION THAT A BOOLEAN EQUATION HAVE A UNIQUE SOLUTION.*

By B. A. BERNSTEIN.

Consider the general Boolean equation

(i) $f(x, y, \cdots, t) \equiv Axy \cdots t + Bxy \cdots t' + \cdots + Lx'y' \cdots t' = 0,$†

involving n unknowns $x, y, \cdots, t$ and having $A, B, \cdots, L$ for its 2^n discriminants. Consider also the equations

(ii) $$x = a, \quad y = b, \cdots, \quad t = k,$$

and the equation

(iii) $\phi(x, y, \cdots, t) \equiv a'x + ax' + b'y + by' + \cdots + k't + kt' = 0.$

Whitehead ‡ has shown that the necessary and sufficient condition that (i) have the unique solution (ii) is that (i) be of the form (iii). Professor Whitehead's proof, carried out only for an equation in two unknowns, consists in showing that the necessary and sufficient condition that (i) have the unique solution (ii) is relation

(iv) $$AB \cdots L + A'B' + A'C' + \cdots + K'L' = 0,$$

which relation he proved earlier to be a necessary and sufficient condition that f of (i) be of the form ϕ of (iii). But Professor Whitehead's proofs require a good bit of calculation, especially for the general case of n variables.§ The object of this note is: (1) to offer a very simple proof of the fact that an equation of form (iii) has the solution (ii) and conversely; (2) to offer a very simple proof of the fact that (iv) is the condition that f of (i) is of the form ϕ of (iii); and (3) to call attention to the simple geometry underlying condition (iv).

* Presented to the Society, April 11, 1931.

† The usual Boole-Schröder notation is used, except that a' denotes the negative of a.

‡ A. N. Whitehead, "Memoir on the Algebra of Symbolic Logic," *American Journal of Mathematics*, Vol. 23 (1901), pp. 140-150.

§ To obtain (iv) as the necessary condition that (i) have a unique solution Whitehead eliminates from (i) all the variables except one, in turn, then gets the conditions that the resulting equations have unique solutions, and then combines these conditions into a single condition. To obtain (iv) as the condition that f of (i) be of the form ϕ of (iii), Professor Whitehead develops f normally with respect to the variables, then identifies corresponding discriminants, and then simplifies.

1. A very simple proof of the fact that an equation of form (iii) has the solution (ii) and conversely, consists in merely observing that equation (iii) and the system of equations (ii) are each equivalent to the system

$$\text{(v)} \qquad a'x + ax' = 0, \quad b'y + by' = 0, \cdots, \quad k't + kt' = 0.$$

It will be noted that this proof does not use (iv) at all.

2. To obtain (iv) very simply as the necessary and sufficient condition that f of (i) be of the form ϕ of (iii), observe that the discriminants of ϕ are

$$\phi(1, 1, \cdots, 1) = a' + b' + \cdots + k', \quad \phi(1, 1, \cdots, 0) = a' + b' + \cdots + k,$$
$$\phi(0, 0, \cdots, 0) = a + b + \cdots + k.$$

Hence, f will be identical with ϕ when and only when

$$A = a' + b' + \cdots + k', \quad B = a' + b' + \cdots + k, \cdots, \quad L = a + b + \cdots + k,$$

or

$$\text{(vi)} \qquad A' = ab \cdots k, \quad B' = ab \cdots k', \cdots, \quad L' = a'b' \cdots k'.$$

But the right-hand members of (vi) are seen to be the *constituents* in the normal development of 1 with respect to $a, b, \cdots, k$. Hence,

$$A' + B' + \cdots + L' = 1, \quad A'B' = 0, \quad A'C' = 0, \cdots, \quad K'L' = 0,$$

or

$$\text{(iv)} \qquad AB \cdots L + A'B' + A'C' + \cdots + K'L' = 0.$$

Of course, since (ii) is equivalent to (iii), we have that (iv) is also the condition that (i) have the solution (ii).

3. The simple geometry underlying condition (iv) is seen from (vi). The necessary and sufficient condition that (i) have the solution (ii), or that f of (i) be of the form ϕ of (iii), is, geometrically, that the regions representing the negatives of the discriminants of f be the regions corresponding to the 2^n constituents in the normal development of 1 with respect to $a, b, \cdots, k$ of ϕ. The geometry for the general equation (i) is thus a very simple extension of the geometry for the equation in one unknown, $a'x + ax' = 0$.

UNIVERSITY OF CALIFORNIA,
BERKELEY, CALIFORNIA.

ON THE EQUIVALENCE OF THE DECOMPOSITIONS OF AN ALGEBRA WITH RESPECT TO A PRINCIPAL IDEMPOTENT.

By Laurens Earle Bush.

1. *Introduction.* In the thesis of the author entitled *Some Properties of Algebras without Moduli* * the results obtained were apparently dependent upon the decomposition of the algebra with respect to a definitely selected principal idempotent. This naturally raises a question as to the extent to which this dependence is actual. The purpose of this paper is, at least partially, to answer this question.

2. *Decomposition of an algebra with respect to a principal idempotent.* Let $\mathfrak{A}$ be a non-nilpotent algebra over an infinite field $\mathfrak{F}$. $\mathfrak{A}$ possesses at least one principal idempotent element u.† Let $\mathfrak{B}$ be the set of all elements x of $\mathfrak{A}$ such that $uxu = 0$, and $\mathcal{F}$ the set of all elements x of $\mathfrak{A}$ such that $ux = xu = 0$. Then

$$\mathfrak{A} = u\mathfrak{A}u + u\mathfrak{B} + \mathfrak{B}u + \mathcal{F},\ddagger$$

where the linear systems $u\mathfrak{A}u$, $u\mathfrak{B}$, $\mathfrak{B}u$, and $\mathcal{F}$ are supplementary in $\mathfrak{A}$. We have further that $u\mathfrak{A}u = \mathfrak{S} + \bar{\mathfrak{N}}$;§ where $\mathfrak{S}$ is a semi-simple subalgebra of $\mathfrak{A}$, not a zero algebra of order one, having the modulus u and $\bar{\mathfrak{N}}$ is the radical (maximal nilpotent invariant subalgebra) of $u\mathfrak{A}u$, and

$$\bar{\mathfrak{N}} + u\mathfrak{B} + \mathfrak{B}u + \mathcal{F} = \mathfrak{N},$$

where $\mathfrak{N}$ is the radical of $\mathfrak{A}$.

It is easily shown that the elements of the systems $\mathfrak{S}$, $\bar{\mathfrak{N}}$, $u\mathfrak{B}$, $\mathfrak{B}u$, and $\mathcal{F}$ form products according to the following table.

	$\mathfrak{S}$	$\bar{\mathfrak{N}}$	$u\mathfrak{B}$	$\mathfrak{B}u$	$\mathcal{F}$
$\mathfrak{S}$	$\mathfrak{S}$	$\bar{\mathfrak{N}}$	$u\mathfrak{B}$	0	0
$\bar{\mathfrak{N}}$	$\bar{\mathfrak{N}}$	$\bar{\mathfrak{N}}$	$u\mathfrak{B}$	0	0
$u\mathfrak{B}$	0	0	0	$\bar{\mathfrak{N}}$	$u\mathfrak{B}$
$\mathfrak{B}u$	$\mathfrak{B}u$	$\mathfrak{B}u$	$\mathcal{F}$	0	0
$\mathcal{F}$	0	0	0	$\mathfrak{B}u$	$\mathcal{F}$

* *The Ohio State University, Abstracts of Doctors' Dissertations*, No. 7 (1931).

† L. E. Dickson, *Algebren und ihre Zahlentheorie*, Zurich, 1927, p. 100.

‡ J. H. M. Wedderburn, *Proceedings of the London Mathematical Society*, Vol. 6, 2nd series (1907-8), pp. 91-93, and Dickson, *loc. cit.*, pp. 98-99.

§ Dickson, *loc. cit.*, p. 136.

For example, if x and y are respectively in $u\mathfrak{B}$ and $\mathfrak{B}u$, xy is in $\bar{\mathfrak{N}}$, or if x and y are respectively in $\bar{\mathfrak{N}}$ and $\mathfrak{F}$, $xy = 0$.

3. *Some theorems on the principal idempotents of an algebra.* If u_1 and u_2 are principal idempotents of $\mathfrak{A}$, $u_1 - u_2$ is properly nilpotent in $\mathfrak{A}$.* We have conversely

THEOREM 3. 1. *If $u' = u + n$ is idempotent, where u is a principal idempotent of $\mathfrak{A}$ and n is properly nilpotent in $\mathfrak{A}$, u' is a principal idempotent of $\mathfrak{A}$.*

For, let $e = x_1 + x_2$ be any idempotent, where x_1 is in that determination of $\mathfrak{S}$ for which u is a modulus and x_2 is in $\mathfrak{N}$. Then

$$u'e = (u + n)(x_1 + x_2) = x_1 + nx_1 + ux_2 + nx_2.$$

Since e is idempotent, $x_1 \neq 0$. But x_1 is in $\mathfrak{S}$, while, on account of the invariance of $\mathfrak{N}$, $nx_1 + ux_2 + nx_2$ is in $\mathfrak{N}$. Hence $u'e \neq 0$, and u' is a principal idempotent.

THEOREM 3. 2. *If u is a principal idempotent of $\mathfrak{A}$ and x is any element of $u\mathfrak{B}$ or $\mathfrak{B}u$, $u + x$ is a principal idempotent of $\mathfrak{A}$.*

For, $(u + x)^2 = u + x$ and $u + x$ is a principal idempotent by theorem 3. 1.

THEOREM 3. 3. *If u is a principal idempotent of $\mathfrak{A}$ and $\mathfrak{B}u\mathfrak{B} = 0$, then $e = u - x_2x_3 + x_2 + x_3$, where x_2 is in $u\mathfrak{B}$ and x_3 in $\mathfrak{B}u$, is a principal idempotent of $\mathfrak{A}$, and every principal idempotent of $\mathfrak{A}$ is of this form.*

For, since $\mathfrak{B}u\mathfrak{B} = 0$, $x_3x_2 = 0$, and therefore $e^2 = e$. By theorem 3. 1, e is a principal idempotent of $\mathfrak{A}$. Conversely, suppose e is any principal idempotent of $\mathfrak{A}$. Then $e - u$ is in $\mathfrak{N}$ by the remark at the beginning of this section, and we may write $e = u + x_1 + x_2 + x_3 + x_4$, where x_1, x_2, x_3, and x_4 are respectively in $\bar{\mathfrak{N}}$, $u\mathfrak{B}$, $\mathfrak{B}u$, and $\mathfrak{F}$. Since $e^2 = e$, we have

$$(x_1 + x_1^2 + x_2x_3) + (x_1x_2 + x_2x_4) + (x_3x_1 + x_4x_3) + (x_4^2 - x_4) = 0.$$

Since $\bar{\mathfrak{N}}$, $u\mathfrak{B}$, $\mathfrak{B}u$, and $\mathfrak{F}$ are supplementary in their sum, the terms in each of these systems must vanish separately, i. e.

$$x_1 + x_1^2 + x_2x_3 = 0, \tag{1}$$
$$x_1x_2 + x_2x_4 = 0, \tag{2}$$
$$x_3x_1 + x_4x_3 = 0, \tag{3}$$
$$x_4^2 - x_4 = 0. \tag{4}$$

* Wedderburn, *loc. cit.*, p. 92.

Since x_4 is nilpotent or zero, (4) requires that $x_4 = 0$. By (2), $x_1x_2 = 0$. Multiplying (1) on the left by x_1 and x_1^2 in turn and subtracting the resulting equations, we have, after transposing terms, $x_1^2 = x_1^4$. Since x_1^2 is nilpotent or zero, this requires that $x_1^2 = 0$. Then, by (1), $x_1 = -x_2x_3$. Hence $e = u - x_2x_3 + x_2 + x_3$.

In a similar manner we may prove

THEOREM 3.4. *If u is a principal idempotent of $\mathfrak{A}$ and $u\mathfrak{B}^2u = 0$ (i. e., if $\mathfrak{B}$ is an algebra), then $e = u + x_2 + x_3 + x_3x_2$, where x_2 is in $u\mathfrak{B}$ and x_3 in $\mathfrak{B}u$, is a principal idempotent of $\mathfrak{A}$, and every principal idempotent of $\mathfrak{A}$ is of this form.*

It is evident that $\mathfrak{A}$ and $\mathfrak{A}^r$ have the same set of idempotents, and therefore also of principal idempotents. If $\mathfrak{A}$ is of index r, $\mathfrak{A}$ therefore has a unique principal idempotent if and only if $\mathfrak{A}^r$ has. If $\mathfrak{A}^r$ has a modulus, it, and consequently $\mathfrak{A}$, has a unique principal idempotent.* If, conversely, $\mathfrak{A}$ has a unique principal idempotent u, by theorem 3.2, $u\mathfrak{B} = \mathfrak{B}u = 0$. Then $\mathfrak{A} = u\mathfrak{A}u + \mathfrak{F}$ and

$$\mathfrak{A}^r = u\mathfrak{A}u + \mathfrak{F}^r = \mathfrak{A}^{r+1} = u\mathfrak{A}u + \mathfrak{F}^{r+1}.$$

Hence, since $\mathfrak{F}$ is nilpotent and $\mathfrak{F}^r = \mathfrak{F}^{r+1}$, we have $\mathfrak{F}^r = \mathfrak{F}^{r+1} = 0$, and $\mathfrak{A}^r = u\mathfrak{A}u$ has the modulus u. Thus we have

COROLLARY 3.41. *A non-nilpotent algebra $\mathfrak{A}$ of index r has a unique principal idempotent if and only if $\mathfrak{A}^r$ has a modulus.*

4. *Equivalence of decompositions.* Let u_1 and u_2 be two different principal idempotents of $\mathfrak{A}$, and let the decompositions of $\mathfrak{A}$ with respect to u_1 and u_2 be respectively

$$\mathfrak{A} = \mathfrak{S}_1 + \bar{\mathfrak{R}}_1 + u_1\mathfrak{B}_1 + \mathfrak{B}_1u_1 + \mathfrak{F}_1,$$

and

$$\mathfrak{A} = \mathfrak{S}_2 + \bar{\mathfrak{R}}_2 + u_2\mathfrak{B}_2 + \mathfrak{B}_2u_2 + \mathfrak{F}_2.$$

These two decompositions will be called *equivalent* if it is possible to associate with every element x of $\mathfrak{A}$ a unique element $\bar{x}$ of $\mathfrak{A}$ in such a manner that

(1) Every element of $\mathfrak{A}$, considered as an $\bar{x}$, is associated with a unique element x;

(2) The association is isomorphic under addition, subtraction, and both types of multiplication;

(3) According as x is in $\mathfrak{S}_1$, $\bar{\mathfrak{R}}_1$, $u_1\mathfrak{B}_1$, $\mathfrak{B}_1u_1$, or $\mathfrak{F}_1$, the associated

---

* Dickson, *loc. cit.*, p. 100.

element $\bar{x}$ is in the corresponding system $\mathfrak{S}_2$, $\bar{\mathfrak{N}}_2$, $u_2\mathfrak{B}_2$, $\mathfrak{B}_2u_2$, or $\mathfrak{F}_2$ respectively.

It is evident from the definition that equivalence of decompositions is a transitive relation.

THEOREM 4.1. *A necessary and sufficient condition that the decompositions of $\mathfrak{A}$ with respect to u_1 and u_2 be equivalent is that it be possible to associate with every element x of $\mathfrak{A}$ a unique element $\bar{x}$ of $\mathfrak{A}$ in such a manner that* (1) *and* (2) *are satisfied, and that* $\bar{u}_1 = u_2$.

For, suppose (1), (2), and (3) are satisfied. If x is in $\mathfrak{S}_1$, $\bar{x}$ is in $\mathfrak{S}_2$. In particular, $\bar{u}_1$ is in $\mathfrak{S}_2$. For every x in $\mathfrak{S}_1$, $u_1x = x$ and $xu_1 = x$, and therefore for every $\bar{x}$ in $\mathfrak{S}_2$, $\bar{u}_1\bar{x} = \bar{x}$ and $\bar{x}\bar{u}_1 = \bar{x}$. That is $\bar{u}_1$ is a modulus for $\mathfrak{S}_2$, and since u_2 is the only modulus for $\mathfrak{S}_2$, $\bar{u}_1 = u_2$.

Conversely, suppose that (1) and (2) are satisfied and that $\bar{u}_1 = u_2$. Let $e_1 = u_1, e_2, e_3, \cdots, e_n$ be a basis for $\mathfrak{A}$, such that $e_1, e_2, \cdots, e_{\rho_1}$ is a basis for $\mathfrak{S}_1$, $e_{\rho_1+1}, e_{\rho_1+2}, \cdots, e_{\rho_2}$ a basis for $\bar{\mathfrak{N}}_1$, $e_{\rho_2+1}, e_{\rho_2+2}, \cdots, e_{\rho_3}$ a basis for $u_1\mathfrak{B}_1$, $e_{\rho_3+1}, e_{\rho_3+2}, \cdots, e_{\rho_4}$ a basis for $\mathfrak{B}_1u_1$, and $e_{\rho_4+1}, e_{\rho_4+2}, \cdots, e_n$ a basis for $\mathfrak{F}_1$. On account of (2), the elements $\bar{e}_1 = u_2, \bar{e}_2, \bar{e}_3, \cdots, \bar{e}_n$, which correspond to $e_1, e_2, \cdots, e_n$ respectively, are linearly independent with respect to $\mathfrak{F}$, and can therefore be selected as a new basis for $\mathfrak{A}$. The multiplication constants for the new basis will be identical with those for the old, and it follows that in the decomposition with respect to u_2, $\bar{e}_1, \bar{e}_2, \cdots, \bar{e}_{\rho_1}$ will form a basis for some determination of $\mathfrak{S}_2$, $\bar{e}_{\rho_1+1}, \bar{e}_{\rho_1+2}, \cdots, \bar{e}_{\rho_2}$ a basis for $\bar{\mathfrak{N}}_2$, $\bar{e}_{\rho_2+1}, \bar{e}_{\rho_2+2}, \cdots, \bar{e}_{\rho_3}$ a basis for $u_2\mathfrak{B}_2$, $\bar{e}_{\rho_3+1}, \bar{e}_{\rho_3+2}, \cdots, \bar{e}_{\rho_4}$ a basis for $\mathfrak{B}_2u_2$, and $\bar{e}_{\rho_4+1}, \bar{e}_{\rho_4+2}, \cdots, \bar{e}_n$ a basis for $\mathfrak{F}_2$. Since if $x = \sum_{i=1}^{n} x_ie_i$, it follows that $\bar{x} = \sum_{i=1}^{n} x_i\bar{e}_i$, (3) is satisfied.

By theorem 3.2, if u is a principal idempotent of $\mathfrak{A}$ and x is in $u\mathfrak{B}$ or $\mathfrak{B}u$, then $u + x$ is a principal idempotent of $\mathfrak{A}$. We shall prove the following theorem regarding the set of principal idempotents obtained in this manner.

THEOREM 4.2. *If u is a principal idempotent of $\mathfrak{A}$ and x is in $u\mathfrak{B}$ or $\mathfrak{B}u$, the decomposition of $\mathfrak{A}$ with respect to $e = u + x$ is equivalent to the decomposition of $\mathfrak{A}$ with respect to u.*

In order to prove this we shall set up a correspondence satisfying the conditions of theorem 4.1. Let us suppose that x is in $u\mathfrak{B}$ and let $\mathfrak{A} = u\mathfrak{A}u + u\mathfrak{B} + \mathfrak{B}u + \mathfrak{F} = e\mathfrak{A}e + e\bar{\mathfrak{B}} + \bar{\mathfrak{B}}e + \mathfrak{F}$. Let $y = y_1 + y_2 + y_3 + y_4$, where y_1, y_2, y_3, and y_4 are respectively in $u\mathfrak{A}u$, $u\mathfrak{B}$, $\mathfrak{B}u$, and $\mathfrak{F}$, be any element of $\mathfrak{A}$. To y we let correspond the element

$$\bar{y} = y + y_1x + y_3x - xy_3 - xy_4 - xy_3x,$$

which is uniquely determined by y. Conversely, if $\bar{y} = \bar{y}_1 + \bar{y}_2 + \bar{y}_3 + \bar{y}_4$, where $\bar{y}_1$, $\bar{y}_2$, $\bar{y}_3$, and $\bar{y}_4$ are respectively in $u\mathfrak{A}u$, $u\mathfrak{B}$, $\mathfrak{B}u$, and $\mathfrak{F}$, is any element of $\mathfrak{A}$, it is the correspondent by the above formula of the uniquely determined element $y = \bar{y} + x\bar{y}_3 + x\bar{y}_4 - \bar{y}_1 x - \bar{y}_3 x - x\bar{y}_3 x$. Direct computation shows that this correspondence is isomorphic under addition, subtraction, and both kinds of multiplication. Also $\bar{u} = u + ux = u + x = e$. Thus the conditions of theorem 4. 1 are satisfied. The proof is similar when x is in $\mathfrak{B}u$.

As an illustration of theorem 4. 2, consider the algebra defined by the following table.

	e_1	e_2	e_3	e_4	e_5
e_1	e_1	e_2	e_3	0	0
e_2	e_2	0	0	0	0
e_3	0	0	0	e_2	0
e_4	e_4	0	0	0	0
e_5	0	0	0	0	0

The element e_1 is a principal idempotent and in the decomposition with respect to it we have e_1, e_2, e_3, e_4, and e_5 respectively in $\mathfrak{S}$, $\bar{\mathfrak{N}}$, $u\mathfrak{B}$, $\mathfrak{B}u$, and $\mathfrak{F}$. The element $e_1 + e_3$ is a principal idempotent of the form required by theorem 4. 2. In the correspondence set up by the proof of this theorem, we have that e_1, e_2, e_3, e_4, and e_5 correspond respectively to $e_1 + e_3$, e_2, e_3, $e_4 - e_2$, and e_5. If we select the latter five elements as a new basis, calling them respectively e'_1, e'_2, e'_3, e'_4, and e'_5, we obtain a new multiplication table which is identical with the original one with e_i $(i = 1, 2, 3, 4, 5)$ replaced by e'_i.

THEOREM 4. 3. *If u is a principal idempotent of $\mathfrak{A}$ and $\mathfrak{B}u\mathfrak{B} = 0$, the decomposition of $\mathfrak{A}$ with respect to any principal idempotent is equivalent to that with respect to u.*

For, let e be any principal idempotent of $\mathfrak{A}$. By theorem 3. 3, $e = u - x_2x_3 + x_2 + x_3$, where x_2 and x_3 are respectively in $u\mathfrak{B}$ and $\mathfrak{B}u$. By theorem 3. 2, $\bar{e} = u + x_2$ is a principal idempotent of $\mathfrak{A}$, and by theorem 4. 2, the decomposition of $\mathfrak{A}$ with respect to $\bar{e}$ is equivalent to that with respect to u. Let $\mathfrak{A} = \bar{e}\mathfrak{A}\bar{e} + \bar{e}\bar{\mathfrak{B}} + \bar{\mathfrak{B}}\bar{e} + \bar{\mathfrak{F}}$. We have $e = \bar{e} + y$, where $y = -x_2x_3 + x_3$. Let $y = y_1 + y_2 + y_3 + y_4$, where y_1, y_2, y_3, and y_4 are respectively in $\bar{e}\mathfrak{A}\bar{e}$, $\bar{e}\bar{\mathfrak{B}}$, $\bar{\mathfrak{B}}\bar{e}$, and $\bar{\mathfrak{F}}$. Then $\bar{e}y = y_1 + y_2$. But

$$\bar{e}y = (u + x_2)(-x_2x_3 + x_3) = 0.$$

Hence $y_1 + y_2 = 0$, and $y = y_3 + y_4$. Then $e = \bar{e} + y_3 + y_4$ is idempotent, which demands that $y_4y_3 + y_4^2 - y_4 = 0$. But y_4y_3 and $y_4^2 - y_4$ are respectively in $\bar{\mathfrak{B}}\bar{e}$ and $\bar{\mathfrak{F}}$ and must therefore vanish separately. Since y_4 is nil-

potent or zero, $y_4^2 = y_4$ requires that $y_4 = 0$, i. e., $e = \bar{e} + y_3$, where y_3 is in $\mathfrak{B}\bar{e}$, and, by theorem 4. 2, the decomposition of $\mathfrak{A}$ with respect to e is equivalent to that with respect to $\bar{e}$, and consequently to that with respect to u.

By a similar proof we may establish the following theorem.

THEOREM 4. 4. *If u is a principal idempotent of $\mathfrak{A}$ and $u\mathfrak{B}^2u = 0$, the decomposition of $\mathfrak{A}$ with respect to any principal idempotent is equivalent to that with respect to u.*

5. *Modulus and semi-modulus.*

THEOREM 5. 1. *If u is a principal idempotent of $\mathfrak{A}$, $\mathfrak{A}$ possesses a modulus if and only if $\mathfrak{B} = 0$ in the decomposition of $\mathfrak{A}$ with respect to u.*

For, if $\mathfrak{B} = 0$, u is evidently a modulus for $\mathfrak{A}$. But if $\mathfrak{A}$ possesses a modulus u, it is the only principal idempotent of $\mathfrak{A}$ and $\mathfrak{A} = u\mathfrak{A}u$.

If $\mathfrak{A}$ possesses an element u such that $ux = x$ $(xu = x)$ for every x in $\mathfrak{A}$, but $xu \neq x$ $(ux \neq x)$ for some x in $\mathfrak{A}$, we shall call u a *left* (*right*) *hand semi-modulus of* $\mathfrak{A}$.

THEOREM 5. 2. *A semi-modulus of $\mathfrak{A}$ is a principal idempotent of $\mathfrak{A}$.*

For, if u is a semi-modulus of A and e any idempotent of $\mathfrak{A}$, either $ue = e \neq 0$ or $eu = e \neq 0$.

THEOREM 5. 3. *If $\mathfrak{A}$ has a semi-modulus, every principal idempotent of $\mathfrak{A}$ is a semi-modulus of the same type.*

For, let u be a left-hand semi-modulus of $\mathfrak{A}$. Then

$$\mathfrak{B}u + \mathfrak{J} = u(\mathfrak{B}u + \mathfrak{J}) = 0 \quad \text{and} \quad \mathfrak{A} = u\mathfrak{A}u + u\mathfrak{B}.$$

By theorem 3. 3, every principal idempotent is of the form $u + x_2$, where x_2 is in $u\mathfrak{B}$. If y is any element of $\mathfrak{A}$, we have

$$(u + x_2)y = (u + x_2)uy = uy = y.$$

A similar proof holds if u is a right-hand semi-modulus.

COROLLARY 5. 31. *If $\mathfrak{A}$ has a semi-modulus it has an infinite number of them.*

For, if u is a left-hand semi-modulus, $u\mathfrak{B} \neq 0$ and contains an infinite number of elements x. For each of these $u + x$ is a left-hand semi-modulus.

COROLLARY 5. 32. *If $\mathfrak{A}$ has a semi-modulus, the decompositions of $\mathfrak{A}$ with respect to principal idempotents are all equivalent.*

THE OHIO STATE UNIVERSITY,
COLUMBUS, OHIO.

Sec. No..........
Book No........
LIBRARY.

HYPERELLIPTIC FUNCTIONS AND IRRATIONAL BINARY INVARIANTS. I.

By Arthur B. Coble.

Introduction. Much of the theory of the theta functions, both in the hyperelliptic and in the general case, is based on the theorem that all functions of given order and characteristic can be expressed linearly in terms of a definite number, k, of such functions. It thus becomes important to have at least one linearly independent set of k functions in terms of which all functions of the given order and characteristic can be linearly expressed. Nevertheless many applications of great interest are based on the existence of quite special three-term linear relations connecting particular functions (cf. [4], p. 441). These special relations must be consequences of the more general linear relations connecting $k+1$ functions. Yet the special relations can be obtained from the more general relations only by some knowledge of the particular behavior of the coefficients in these relations.

Usually the coefficients of the theta relations are values of the theta functions for particular values of the arguments; or, if one prefers, they are the values of given functions with given characteristics for the zero values of the arguments. These coefficients depend only on the moduli of the theta functions. It is therefore desirable, as Wirtinger has pointed out (cf. [5], p. 69), to consider theta relations as relations connecting functions both of the moduli and of the arguments with coefficients in the domain of the ordinary number system.

A similar situation exists in the theory of invariants. Four quadratic covariants of a system of binary forms are necessarily linearly dependent when considered as polynomials in the binary variables alone. Usually however the coefficients of the linear relation are invariants, and in such case the four quadratics are independent in the numerical domain.

The hyperelliptic thetas, defined by a binary form of order $2p+2$, can be exhibited very satisfactorily, in the case $p=2$, by beginning with a numerical domain. It is the object of the present series of papers to consider the possible extensions of the situation for $p=2$ to higher values of p. We outline briefly here the case, $p-2$. Having chosen in a linear space, $S_3(y)$, the five points which determine a coördinate system, and which therefore have only numerical coördinates, the system (∞^4) of quadrics on these five points

maps $S_3(y)$ upon the points of a locus $M_3^3(x)$ in S_4 whose equation may be put into the form, $\sum_{i=1}^{i=6} x_i^3 = 0$ $\{\sum_{i=1}^{i=6} x_i \equiv 0\}$. A binary sextic with roots $t_1, \cdots, t_6$ and differences (ik) has 15 irrational linear invariants (A) of the form $(12)(34)(56)$ of which only five are linearly independent in the domain $R(1)$. Five that are independent map the ∞^3 projectively distinct binary sextics upon the points of $M_3^3(x)$. The sextic also has ten irrational invariants (B) of the second degree and of the form $(12)(13)(23)\cdot(45)(46)(56)$, of which again five are linearly independent. Five which are linearly independent, and which are properly chosen with respect to the five invariants (A), map the ∞^3 binary sextics upon the spaces ξ of an envelope, $M_3^4(\xi)$, of class four, which is the dual form of $M_3^3(x)$ in such wise that x, ξ are point, and tangent space at the point, of $M_3^3(x)$. Thus the moduli of the theta functions are determined algebraically by the choice of point x, or space $S_3(\xi)$, of the "modular manifolds" $M_3^3(x)$, $M_3^4(\xi)$, which themselves have no absolute invariants.

In § 1 we prove, for the general $(2p+2)$-ic, that the system of invariants (A), and the system of invariants (B), have the same dimension, and therefore define modular manifolds, $M_{2p-1}(x)$, $M_{2p-1}(\xi)$ in the same linear space. In § 2 $M_{2p-1}(x)$ is discussed as the map of a linear space, $S_{2p-1}(y)$. The projective construction of a point x of $M_{2p-1}(x)$, determined by a $(2p+2)$-ic with ordered roots, is given in terms of the positions of certain "median" points. For $p=2$, these median points are the ten nodes of $M_3^3(x)$. In §§ 3, 4 the linear relations connecting these median points, and the construction mentioned, are discussed analytically. In § 5 we prove that the x, ξ are dually related, and that x, ξ defined by the same $(2p+2)$-ic are incident.

Returning again to the case $p=2$, when x is chosen on $M_3^3(x)$, the enveloping cone of $M_3^3(x)$ from x is a Kummer cone K whose section by an S_3 not on x is a Kummer surface; and the polar quadric of x as to $M_3^3(x)$ cuts $M_3^3(x)$ in an M_2^6 which is the map from $S_3(y)$ of a Weddle quartic surface. The birational transformation from Weddle to Kummer surface is immediate. If y is a point on the Weddle which maps into x' on M_3^3, the tangent S_3 of M_3^3 at x' is on x, and therefore on K, and contributes, in the section, a tangent plane of the Kummer surface. Hence the domain (∞^5) of the moduli, and of the variables, of the theta functions $p=2$ is represented by pairs of points x, x' on M_3^3 such that the tangent S_3 at x' is on x.

The extension of the Kummer surface to a Kummer manifold K_p of dimension p for the general theta functions of genus p has been developed by Wirtinger. In the hyperelliptic case K_p still exists, but doubtless has many interesting special properties. The author has recently developed the corre-

sponding hyperelliptic extension of the Weddle W_p, and has given the mapping by means of which these manifolds are birationally related. In § 6 some applications to W_3 and to W_p are given. But the question still remains as to whether the K_p and W_p may not be projectively attached to a point of a modular manifold M_{2p-1} in such wise that their birational connection is as evident as in the case $p = 2$.

The author has also proved ([3], pp. 144-148) that the configuration of nodes and tropes of the Kummer, as well as algebraic parametric representations of both the Kummer and the Weddle, can be given in quite simple form in the case $p = 2$ in terms of irrational invariants and covariants of the sextic. We shall examine later the possibility of extending these representations to higher values of p.

As at present planned, part II of this series will be devoted to a more careful examination of the particular case $p = 3$; and part III to a discussion of the loci on $M_{2p-1}(x)$ which correspond to $(2p+2)$-ics with multiple roots of any description.

1. The number of linearly independent linear irrational covariants of weight k for a binary form of order $j = 2k + l$: Let $F = (\alpha t)^j$ be a binary form of order j whose leading coefficient we take for convenience as unity. If the roots of $F = 0$ in nonhomogeneous form are $t = t_1, t_2, \cdots, t_j$, let

$$(1) \qquad (ij) = t_i - t_j, \qquad (it) = (t_i - t).$$

If then we select from the j roots k pairs, and form the product of differences,

$$(2) \qquad \pi = (12)(34)\cdots(2k-1, 2k)(2k+1, t)\cdots(2k+l, t);$$

this product is an irrational covariant of F of *degree* 1, *weight* k, and *order* l ([1], I, § 3).

The number of distinct covariants of this type is

$$(3) \qquad \binom{j}{l}\frac{(2k)!}{k!\,2^k} = \frac{j!}{k!\,l!\,2^k} \qquad (j = 2k + l).$$

However, by using the binary identities, a particular product may be expressed in terms of others with coefficients which are free of the roots, i. e., these covariants are related linearly and with integer coefficients in many ways. We ask for the number,

$$(4) \qquad n_{k,l} \qquad (j = 2k + l),$$

which are linearly independent in the ring of integers.

The products π which vanish when $t = t_j$ are those which contain the

factor (jt). The coefficients of (jt) in these products lie in the linear system of irrational linear covariants of weight k and order $l-1$ belonging to the form $F/(jt)$. Thus the number of linearly independent coefficients of (jt) is $n_{k,l-1}$. If π does not contain t_j in the form (jt), π must have the form

$$(1j)(2t)(3t)\cdots(l+1,t)\cdot\pi'.$$

Since $(1j)(2t)=(12)(jt)+(1t)(2j)$, it is clear that, to within products containing (jt), i.e., to within the linear system already accounted for, the terms,

$$\begin{array}{l}(1j)(2t)(3t)\ \cdots\ (l+1,t)\ \cdot\pi',\\ (1t)(2j)(3t)\ \cdots\ (l+1,t)\ \cdot\pi',\\ (1t)(2t)(3j)\ \cdots\ (l+1,t)\ \cdot\pi',\\ \cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\\ (1t)(2t)(3t)\ \cdots\ (l+1,j)\ \cdot\pi',\end{array}$$

are equal to each other. But the sum of these terms divided by $l+1$ is the polar of t_j as to a covariant of weight $k-1$ and order $l+1$. Since all of these covariants can be expressed in terms of $n_{k-1,l+1}$, we have the recursion formula

$$n_{k,l}=n_{k,l-1}+n_{k-1,l+1} \qquad (l \geqq 1). \tag{5}$$

In the derivation of (5) the existence of factors of the type (jt) was assumed. If none such occur, i.e., if we seek $n_{k,0}$ $(j=2k)$, the number of linearly independent invariants of type $(12)(34)\cdots(j-1,j)$, then, as before, there is a linear system which has the factor (12), and of these $n_{k-1,0}$ are linearly independent. The remaining invariants are of type $(13)(24)\pi'$. This type pairs with $(14)(23)\pi'$. The difference of the two contains the factor (12), and is in the linear system accounted for. The sum is twice the polar of t_1, t_2 as to the quadratic $(3t)(4t)\cdot\pi'$ which lies in a linear system containing $n_{k-2,2}$ linearly independent terms. Hence

$$n_{k,0}=n_{k-1,0}+n_{k-2,2}. \tag{6}$$

If we call $r+s$ the *rank* of the number $n_{r,s}$, we observe that the recursion formulae (5), (6) give the value of $n_{k,l}$ in terms of numbers $n'_{k',l'}$ for which the rank $k'+l'$ is either smaller than $k+l$, or equal to $k+l$ when $l'>l$. Since the maximum value of l is j, and then there is a single covariant F, i.e., $n_{0,j}=1$, the formulae (5), (6) suffice to determine the numbers $n_{k,l}$ completely.

A table of early values calculated from these formulae is:

$$
\begin{aligned}
&n_{0,j} = 1 && (j = 1, 2, \cdots, 10),\\
&n_{1,l} = 1, 2, 3, \cdots, 9 && (l = 0, 1, 2, \cdots, 8),\\
&n_{2,l} = 2, 5, 9, 14, 20, 27, 35 && (l = 0, \cdots, 6),\\
&n_{3,l} = 5, 14, 28, 48, 75 && (l = 0, \cdots, 4),\\
&n_{4,l} = 14, 42, 90 && (l = 0, 1, 2),\\
&n_{5,0} = 42.
\end{aligned}
$$

From an inspection of the table we find that

$$
(7) \qquad n_{k,l} = \binom{2k+l}{k} - \binom{2k+l}{k-1} = \binom{2k+l}{k-1}\frac{l+1}{k}.
$$

It is easy to verify that this value of $n_{k,l}$ satisfies the recursion formulae (5), (6), and therefore it is the value sought.

For the particular case when $l = 0$ and $k = p+1$ the linear covariants become the linear invariants of the binary $(2p+2)$-ic of type (A), i. e., of type $(12)(34)(56)\cdots(2p+1, 2p+2)$. For these it is clear that

(8) *The number of linearly independent irrational linear invariants of the binary* $(2p+2)$*-ic is*

$$
\nu = n_{p+1,0} = \binom{2p+2}{p}\frac{1}{p+1} = \binom{2p+2}{p+1}\frac{1}{p+2}.
$$

Another type of irrational invariant of the binary $(2p+2)$-ic arises from the norm-curve in S_p with canonical equations,

$$
(9) \qquad z_0 = t^p, \; z_1 = t^{p-1}, \; z_2 = t^{p-2}, \cdots, z_p = 1.
$$

The condition that the $(p+1)$ points of this curve which are determined by $t = t_1, \cdots, t_{p+1}$ be on an S_{p-1} (a condition which entails a degeneration of the norm-curve) is

$$
(10) \qquad (1\,2\,3\cdots, p+1) = \begin{vmatrix} t_1{}^p & t_1{}^{p-1} & \cdots & 1 \\ t_2{}^p & t_2{}^{p-1} & \cdots & 1 \\ \cdot & \cdot & \cdot & \cdot \\ t^p_{p+1} & t^{p-1}_{p+1} & \cdots & 1 \end{vmatrix}
= (12)(13)(23)\cdots(p, p+1) = 0.
$$

The product of two such complementary determinants,

$$
(1\,2\,3\cdots, p+1)(p+2, p+3, \cdots, 2p+2),
$$

is an irrational invariant of type (B) of the binary $(2p+2)$-ic, whose degree is p. The invariants of type (B) are connected by linear relations with integer coefficients, which arise from determinant identities. According to [1, I, p. 188(70)]

(11) *The number of linearly independent invariants of type* (B) *is the same as the number of linearly independent invariants of type* (A).

In § 5 we develop a duality between these two types of invariants.

2. The modular variety M_{2p-1} in $S_{\nu-1}$. The invariants (A) for two projectively equivalent binary $(2p+2)$-ics are proportional. If then we select ν linearly independent invariants (A) according to § 1 (8), and set

$$x_i = (i_1 i_2)(i_3 i_4)\cdots(i_{2p+1}\, i_{2p+2}) \qquad (i = 1, \cdots, \nu), \tag{1}$$

the aggregate of projectively distinct and ordered binary $(2p+2)$-ics is mapped upon the points of a variety M_{2p-1} of dimension $2p-1$ in a linear space $S_{\nu-1}$.

The choice of the independent set x_i may be made in many ways. For a particular value of p an available set may be obtained by following out the recurrences of § 1. Thus for $p = 0, 1, 2, 3$ independent sets are

$$\begin{array}{lll}
 & p=0: & (12);\\
 & p=1: & (12)(34),\ (13)(24);\\
 & p=2: & (12)(34)(56),\ (12)(35)(46),\ (14)(25)(36),\\
 & & (13)(24)(56),\ (13)(25)(46);\\
(2) & p=3: & (12)(34)(56)(78),\ (12)(34)(57)(68),\ (12)(36)(47)(58),\\
 & & (13)(24)(56)(78),\ (13)(24)(57)(68),\ (13)(26)(47)(58),\\
 & & (12)(35)(46)(78),\ (12)(35)(47)(68),\ (14)(26)(37)(58),\\
 & & (13)(25)(46)(78),\ (13)(25)(47)(68),\ (15)(26)(37)(48),\\
 & & (14)(25)(36)(78),\ (14)(25)(37)(68).
\end{array}$$

A notable property of M_{2p-1} is:

(3) *The modular variety M_{2p-1} is a rational variety, which is invariant under a collineation group $C_{(2p+2)!}$ which is isomorphic with the symmetric group of order $(2p+2)!$*

For, if the root t_{2p+2} be transformed to $t = \infty$, and at the same time the remaining roots are transformed so that their sum is zero, then these remaining roots, say $y_1, \cdots, y_{2p+1}$, for which

$$y_1 + y_2 + \cdots + y_{2p+1} = 0, \tag{4}$$

are determined only to within a factor of proportionality. Thus the ordered binary $(2p+2)$-ic is represented by the point y in the linear space $S_{2p-1}(y)$, and conversely a point of $S_{2p-1}(y)$ determines the ordered $(2p+2)$-ic to within a projectivity. In the mapping (1) the binary $(2p+2)$-ic, with roots y_i, ∞ furnishes a point

(5) $$x_i = (y_{i_1} - y_{i_2})(y_{i_3} - y_{i_4}) \cdots,$$

the factor $(i_k i_l)$ which contains t_{2p+2} being simply $+1$ or -1 according as i_k or i_l is $2p+2$. Thus x on M_{2p-1} has coördinates which are rational in the coördinates y of a linear S_{2p-1}, and M_{2p-1} is a rational variety. If the roots of the given $(2p+2)$-ic are permuted, the representative point y is transformed by an operation of a Cremona $G_{(2p+2)!}$ in $S_{2p-1}(y)$—the cross-ratio group of Moore [cf. [1], I, § 7 $(k=1)$]. The varieties $x_i = 0$ in (5) make up the simplest linear system of varieties which is invariant under the Cremona $G_{(2p+2)!}$. When $S_{2p-1}(y)$ is mapped by this linear system upon M_{2p-1}, the Cremona group induces a collineation group $C_{(2p+2)!}$ in $S_{\nu-1}$ under which M_{2p-1} is invariant. The equations of these collineations can be obtained from (1) by permuting the roots t, and expressing the resulting products x'_i linearly in terms of the original products x_i.

The $2p+1$ coördinates y, subject to (4), determine in S_{2p-1} a *base* whose $2p+1$ points have typical coördinates

(6) $$-2p,\ 1,\ 1, \cdots, 1.$$

Clearly all of the p factors $(y_{i_k} - y_{i_l})$ vanish at such a basis point except one at most; all except two at most vanish at two basis points; etc., whence

(7) *M_{2p-1} is the map of $S_{2p-1}(y)$ by the linear system of spreads of order p which contain the $\binom{2p+1}{1}$ points of a basis in S_{2p-1} as $(p-1)$-fold points; and which therefore also contain the $\binom{2p+1}{2}$ basal lines $p-2$ times; the $\binom{2p+1}{3}$ basal planes $p-3$ times; $\cdots$; and finally the $\binom{2p+1}{p-1}$ basal S_{p-2}'s simply.*

For, if we set $z_i = y_i - y_{2p+1}$ $(i = 1, \cdots, 2p)$, then $2p$ of the basis points become the reference points, and the last basis point becomes the unit point. The general spread of order p with $(p-1)$-fold points at the reference points is $\Sigma\, a_{12\ldots p} z_1 z_2 \cdots z_p = 0$ with $\binom{2p}{p}$ terms. This will also have a $(p-1)$-fold point at the unit point if its polar of order $p-2$ vanishes. This polar has $\binom{2p}{p-2}$ terms of the form $z_1 z_2 \cdots z_{p-2}$, whence the linear system in (7) contains

(8) $$\nu = \binom{2p}{p} - \binom{2p}{p-2}$$

independent members.

We seek now certain sets of points or sets of linear spaces conjugate under the $C_{(2p+2)!}$ in (3), which are of significance for M_{2p-1}. We begin with a set of $\frac{1}{2}\binom{2p+2}{p+1} = \nu(p+2)/2$ *median points.* These are defined as follows: Consider those $(2p+2)$-ics for which the $p+1$ roots, $t_{p+2}, t_{p+3}, \cdots, t_{2p+2}$

become equal to t. Then all of the invariants (A) either vanish or are equal to $\pm (1t)(2t)\cdots(p+1, t)$. Thus the coördinates x_i in (1) become, on dividing by this factor of proportionality, $x_i = \epsilon_i$, where ϵ_i is $+1, -1$, or 0. If, on the other hand, the complementary set of roots all become equal to s, the same invariants (A) vanish as before, and the non-vanishing ones become equal to $\pm (s, p+2)(s, p+3)\cdots(s, 2p+2)$, i. e., the coördinates x_i are proportional to the same set ϵ_i as before. Hence

(9) *There exists on M_{2p-1} a conjugate set of $\nu(p+2)/2$ median points $P_{12,\ldots,p+1} = P_{p+2,\ldots,p+2}$ which are the maps of those ordered $(2p+2)$-ics for which one set of $p+1$ roots (or the complementary set) are all equal.*

These points are singular points of the mapping since the ∞^{p-1} projectively distinct $(2p+2)$-ics which depend upon the double ratios of $t_1, \cdots, t_{p+1}, t$ all map on the same median point.

Consider now the aggregate of $(2p+2)$-ics for which only p of the roots, say the last p, coincide at t. The invariants (A) which now do not vanish have the form

$$(10) \qquad (12)(3t)(4t)\cdots(p+2, t).$$

Of these, according to § 1 (7), there are $n_{1,p} = p+1$ which are linearly independent, and the $(2p+2)$-ics are mapped upon the points of an S_p. As t changes in (10) the mapping point runs over a norm curve of order p in S_p which for $t = t_1, t_2, \cdots, t_{p+2}$ passes through the $p+2$ median points, $P_{1,p+3,\ldots,2p+2}, P_{2,p+3,\ldots,2p+2}, \cdots, P_{p+2,p+3,\ldots,2p+2}$, in the S_p. For variation of $t_1, \cdots, t_{p+2}$ we obtain the system of ∞^{p-1} norm curves on the base in S_p which consists of these median points. The number of such S_p's is $\binom{2p+2}{p} = (p+1)\nu$. Hence

(11) *There are $\nu(p+1)$ sets of $p+2$ median points whose indices have p common subscripts such that each set lies in an S_p which is contained entirely on M_{2p-1}. The points of such an S_p map binary $(2p+2)$-ics with a p-fold root t. As t varies, the map describes a norm-curve N^p on the $p+2$ median points; and as the remaining roots vary, the N^p runs over a basic system on the $p+2$ points. On each median point there are $2p+2$ such spaces S_p.*

Consider again the aggregate of $(2p+2)$-ics whose last $p-1$ roots coincide at t. The invariants (A) which do not vanish have the form

$$(12) \qquad (12)(34)(5t)\cdots(p+3, t).$$

According to § 1 (7), there are $p(p+3)/2$ of these which are linearly in-

dependent. Such $(2p+2)$-ics are mapped upon the points of a manifold M_{p+1} of dimension $p+1$ in a linear space $S_{-1+p(p+3)/2}$. As t varies in (12) the mapping point runs over a norm-curve N^{p-1} (necessarily in an S_{p-1}), and as t takes in succession the values $t_1, \cdots, t_{p+3}$ the mapping point strikes successively a point of one of $p+3$ of the spaces S_p of (11). Since the N^{p-1} is defined by $p+2$ points, the question arises as to whether an S_{p-1} which meets $p+2$ of the spaces S_{p-1} will also meet the remaining one. We find that there are $\infty^{p^2(p+1)/2}$ S_{p-1}'s in $S_{-1+p(p+3)/2}$, and that it is $p(p-1)/2$ conditions that an S_{p-1} shall meet an S_p. Hence there are ∞^p S_{p-1}'s which meet the given $p+2$ S_p's, one on each point of a given S_p. The N^{p-1} on these $p+2$ intersections passes through the point where the S_{p-1} meets the $(p+3)$-th S_p. Hence

(13) *There are $\binom{2p+2}{p-1}$ sets of $(p+3)$ of the spaces S_p of* (11), *each set lying in a linear space $S_{-1+p(p+3)/2}$. In $S_{-1+p(p+3)/2}$ the S_{p-1}'s which meet $p+2$ of the S_p's are ∞^p in number, one on each generic point of an S_p. Such S_{p-1}'s also meet the $(p+3)$-th S_p, and the $p+3$ intersections of the S_{p-1} with the $p+3$ S_p's lie on a norm-curve N^{p-1}. The locus of these ∞^p N_{p-1}'s is a manifold M_{p+1}, the section of M_{2p-1} by $S_{-1+p(p+3)/2}$, and the map of binary $(2p+2)$-ics with a $(p-1)$-fold root.*

From entirely analogous considerations there follows:

(14) *There exists a set of $\binom{2p+2}{p-k+1}$ linear spaces $\{S\}_k$ $(k=1, \cdots, p-1)$ of dimension $-1+\binom{p+k+1}{k-1}(p-k+2)/k$ which cut the modular manifold M_{2p-1} each in an M_{p+k-1}, the map of projectively distinct binary $(2p+2)$-ics with a $(p-k+1)$-fold root. A space $\{S\}_k$ contains $(p+k+1)$ spaces $\{S\}_{k-1}$. Through each point of M_{p+k-1} there passes a norm-curve N^{p-k+1} which meets the $(p+k+1)$ spaces $\{S\}_{k-1}$ each in one point. The parameters of these $(p+k+1)$ points on N^{p-k+1} are projective to the remaining distinct roots of the binary $(2p+2)$-ic. For the extreme case $k=p$, there is one $\{S\}_p = S_{\nu-1}$ containing M_{2p-1} and $\binom{2p+2}{2}$ spaces $\{S\}_{p-1}$. On each point of M_{2p-1} there are $(2p+2)$ lines each of which crosses $(2p+1)$ of the spaces $\{S\}_{p-1}$.*

The theorems given furnish projective constructions for a generic point on M_{2p-1} in terms of the known median points. The median points themselves are obtained from the reference and unit points by harmonic constructions. For example, to obtain the point on M_5 in S_{13} determined by the binary octavic with ordered roots $t_1, \cdots, t_8$ we proceed as follows. In the $\{S\}_1 = S_3$ determined by $t_6 = t_7 = t_8 = t$ we take an N^3 on the 5 S_0's, or median points, determined by $t_1 = t_6 = t_7 = t_8, \cdots, t_5 = t_6 = t_7 = t_8$

with parameters $t_1, \cdots, t_5$ at these five $\{S\}_0$'s in order, and mark the point on this N^3 with parameter t_6. This is the map of the octavic $t_1, \cdots, t_5, t_6, t_6, t_6$. We construct similarly the map of the octavics $t_1, \cdots, t_6, t_5, t_5; \cdots;$ $t_1, \cdots, t_6, t_1, t_1$. These six points lie on an N^2, a conic, with parameters $t_6, t_5, \cdots, t_1$ respectively. On this conic the point t_7 is the map of the octavic $t_1, \cdots, t_7, t_7$. Similarly the maps of the octavics $t_1, \cdots, t_7, t_6; \cdots;$ $t_1, \cdots, t_7, t_1$ are constructed. These seven points lie on a line with parameters $t_7, t_6, \cdots, t_1$ respectively. On this line the point with parameter t_8 is the required point $t_1, \cdots, t_8$.

The manifold M_{2p-1} has, according to (14), $2p+2$ distinct linear rulings. If the M_{2p-1} is mapped as in (7) from an $S_{2p-1}(y)$, then $2p+1$ of the rulings are the maps of the systems of lines each of which is on one of the $2p+1$ basis points in $S_{2p-1}(y)$, and the remaining ruling is the map of the system of norm-curves N^{2p-1} on the basis. These $2p+2$ systems are permuted symmetrically by Moore's $G_{(2p+2)!}$, and thus on M_{2p-1} the $2p+2$ linear rulings are permuted symmetrically by $C_{(2p+2)!}$.

3. Canonical form of the linear identities connecting the invariants (B). The geometric constructions of § 2 have an algebraic parallel which is developed in the next section. In this the median points are assumed as before. These depend upon the division of the $2p+2$ roots into two sets of $p+1$ each, and thus are in one-to-one correspondence with the invariants (B) of § 1 (10), (11), namely:

$$(1) \qquad (B): \quad (i_1 i_2 \cdots i_{p+1})(j_1 j_2 \cdots j_{p+1}).$$

These invariants are determinant products, and a canonical choice of sign for each, and a similar choice of a set of independent linear relations connecting them, is useful in the next section.

Let one root, say the last, t_{2p+2}, of the $(2p+2)$-ic be isolated, and let this root occupy the last place in the second determinant in (1). The remaining elements in each of the two determinants are then written in the natural order in their respective determinants so as to produce

$$(2) \qquad d_{l_1 \ldots l_p} = \epsilon (k_1 k_2 \cdots k_{p+1})(l_1 l_2 \cdots l_p, 2p+2)$$
$$(k_i < k_j \text{ and } l_i < l_j \text{ if } i < j),$$

where ϵ is $+1$ or -1 according as the permutation

$$k_1 \cdots k_{p+1}\, l_1 \cdots l_p$$

from the natural order is even or odd. Thus the $\binom{2p+1}{p}$ products $d_{l_1 \ldots l_p}$ are determined in sign as soon as the group of subscripts $l_1, \cdots, l_p$ is given.

The linear relations connecting these determinant products are $\binom{2p+2}{p}$ in number. Each relation contains the $p+2$ products which have in one determinant a fixed group of p roots. Thus the relations are of two different types, in the notation introduced above, according as the last root t_{2p+2} is or is not in the fixed group of p roots. For the first type we have

$$(3) \qquad r_{m_1 \ldots m_{p-1}} \equiv d_{k_1 m_1 \ldots m_{p-1}} + \cdots + d_{k_{p+2} m_1 \ldots m_{p-1}} = 0.$$

To prove this we observe merely that in the identity as usually written the signs are alternately $+$ and $-$, since each term arises from the preceding one by an inversion. However, this change of sign due to the inversion has been accounted for by ϵ in the definition (2) of $d_{l_1 \ldots l_p}$.

In the second type we have a fixed group $m_1, \cdots, m_p$ and a residual group $k_1, \cdots, k_{p+1}, 2p+2$. The terms in the identity are

$$d_{m_1 \ldots m_p},\ d_{k_1 \ldots k_p},\ d_{k_1 \ldots k_{p-1} k_{p+1}}, \cdots,\ d_{k_2 \ldots k_{p+1}}.$$

For the same reason as before the signs are like in all terms except perhaps the first. Thus we have to determine ϵ' in

$$\epsilon' d_{m_1 \ldots m_p} + d_{k_1 \ldots k_p}$$

so that this shall be the pair of terms

$$-(k_1 \cdots k_p\, k_{p+1})\,(m_1 \cdots m_p, 2p+2) + (m_1 \cdots m_p\ k_{p+1})\,(k_1 \cdots k_p, 2p+2).$$

Since only the relative signs are to be determined, we may suppose that $k_1 \cdots k_p$ and $m_1, \cdots, m_p$ are already in the natural order. If α transpositions are required to put $k_1 \cdots k_p\, k_{p+1}\, m_1 \cdots m_p$ into the natural order, then $\alpha + p$ transpositions will put $m_1 \cdots m_p\, k_{p+1}\, k_1 \cdots k_p$ into the natural order, since the p transpositions, $(k_1 m_1), \cdots, (k_p m_p)$ change the one into the other. Thus the pair of terms becomes

$$-(-1)^{\alpha}\, d_{m_1 \ldots m_p} + (-1)^{\alpha+p}\, d_{k_1 \ldots k_p},$$

and the second type of linear relation reads as follows:

$$(4) \qquad r_{m_1 \ldots m_p} \equiv (-1)^{p+1}\, d_{m_1 \ldots m_p} + d_{k_1 \ldots k_p} + d_{k_1 \ldots k_{p-1} k_{p+1}} + \cdots + d_{k_2 \ldots k_{p+1}} = 0.$$

We now prove that the identities of this second type (4) are obtained from those of the first type (3). Let the roots other than t_{2p+2} be divided into two sets of p and $p+1$ each, say $i_1, \cdots, i_p$ and $j_1, \cdots, j_{p+1}$. Then there are $\binom{p}{0}$ products of type (a_0) or $d_{i_1 \ldots i_p}$, $\binom{p}{1}\binom{p+1}{1}$ products of type (a_1) or $d_{i_1 \ldots i_{p-1} j_1}$, $\binom{p}{2}\binom{p+1}{2}$ products of type (a_2) or $d_{i_1 \ldots i_{p-2} j_1 j_2}, \cdots$, finally $\binom{p}{p}\binom{p+1}{p}$ products of type (a_p) or $d_{j_1 \ldots j_p}$. Also there are $\binom{p}{1}$ identities

of type (b_1) or $r_{i_1\ldots i_{p-1}}$, each containing 1 term of type (a_0) and $p+1$ terms of type (a_1); $\binom{p}{2}\binom{p+1}{1}$ identities of type (b_2) or $r_{i_1\ldots i_{p-2}j_1}$, each containing 2 terms of type (a_1) and p terms of type (a_2); $\cdots$; finally $\binom{p}{p}\binom{p+1}{p-1}$ identities of type (b_p) or $r_{j_1\ldots j_{p-1}}$, each containing p terms of type (a_{p-1}) and 2 terms of type (a_p). On adding all identities of the type (b_1); all of the type (b_2); etc., we obtain

$$
\begin{aligned}
p\,\Sigma\,(a_0) + \Sigma(a_1) &= 0,\\
(p-1)\,\Sigma\,(a_1) + 2\,\Sigma\,(a_2) &= 0,\\
\cdots\cdots\cdots\cdots\cdots&\\
2\,\Sigma\,(a_{p-2}) + (p-1)\,\Sigma\,(a_{p-1}) &= 0,\\
\Sigma\,(a_{p-1}) + p\,\Sigma\,(a_p) &= 0.
\end{aligned}
$$

The elimination of $\Sigma(a_1),\cdots,\Sigma(a_{p-1})$ from these p equations yields $r_{i_1\ldots i_p} = (-1)^{p+1}\,\Sigma\,(a_0) + \Sigma\,(a_p) = 0$.

The number of determinant products is $\frac{1}{2}\binom{2p+2}{p+1}$, the number of linear identities of type (3) is $\binom{2p+1}{p-1}$. Since $\frac{1}{2}\binom{2p+2}{p+1} - \binom{2p+1}{p-1} = \binom{2p+2}{p+1} - \binom{2p+2}{p+2}$ is the number of linearly independent determinant products or invariants (B) [cf. § 1 (11)], we have shown that

(5) *With the sign of a determinant product determined as in* (2) *the* $\binom{2p+1}{p}$ *products* $d_{l_1\ldots l_p}$ *may be arranged as the elements of a p-way determinant D of order* $2p+1$ *with lines determined by the ordered subscripts* $l_1,\cdots,l_p$. *An element in two or more lines of the same index is to be zero. Then the* $\binom{2p+1}{p-1}$ *independent linear relations of type* (3) *arise by equating the sum of the elements of any line to zero.*

4. Canonical form of the linear identities connecting the median points; equations of the norm-curves of § 2 in terms of the median points. Setting, as in § 2,

(1) $$x_i = (i_1 i_2)(i_3 i_4)\cdots(i_{2p+1}\, i_{2p+2}) \qquad (i=1,\cdots,\nu),$$

let the coördinates ξ_i be contragredient to the point coördinates x in $S_{\nu-1}$. An equation linear in the ξ's is the equation of a point. We prove that the equations of the median points may be affected by such factors of proportionality that their left members satisfy the same system of independent linear relations as the determinant products (B) in § 3 (3).

A particular median point is obtained in (1) when the $p+1$ roots of either of two complementary sets become equal. Let the set containing t_{2p+2} be preferred, and let

(2) $$\xi_{l_1\ldots l_p} = 0$$

be the equation of that median point for which $t_{l_1} = t_{l_2} = \cdots = t_{l_p} = t_{2p+2} = t$. If $t_{j_1}, \cdots, t_{j_{p+1}}$ are the remaining roots, the factor $(j_1 t) \cdots (j_{p+1} t)$ is removed from the coördinates x_i in (1), and the residual factors, $+1$, -1, or 0, are the coefficients of the coördinates ξ in (2). For this canonical choice of the factors of proportionality in the coördinates of the median points their equations satisfy the linear identities,

$$(3) \qquad \xi_{k_1 m_1 \ldots m_{p-1}} + \cdots + \xi_{k_{p+2} m_1 \ldots m_{p-1}} = 0.$$

We have merely to show that a particular coördinate x_i is either 0 at all of the $p+2$ points in (3), or else is $+1$ at one, -1 at one, and 0 at the p remaining points. If x_i contains a factor $(m_i m_j)$, or a factor $(m_i, 2p+2)$, it vanishes at all of the points of (3), and the identity is satisfied. If no two of the $m_1, \cdots, m_{p-1}, 2p+2$ occur in the same factor of x_i, we may suppose that each occurs last in the particular factor which contains it. This amounts to examining $+x_i$, or $-x_i$ as the case may be. Then x_i or $-x_i$ has the form

$$(k'_1 m_1) \cdots (k'_{p-1} m_{p-1}) (k'_p, 2p+2) (k_r k_s)$$

where $k'_1, \cdots, k'_p$ is some permutation of the $k_1, \cdots, k_{p+2}$ other than k_r and k_s. This vanishes when $t_{m_1}, \cdots, t_{m_{p-1}}, t_{2p+2}$ are all equal to t at every point ξ in (3) except at the two points $\xi_{k_r m_1 \ldots m_{p-1}}$ and $\xi_{k_s m_1 \ldots m_{p-1}}$. At the first of these the factor $(k_r k_s)$ becomes $(t k_s)$, and at the second $(k_r t)$. On removing the factors $(k_s t)$, $(k_r t)$ as part of the factor of proportionality to be deleted, the first point has the value -1, and the second the value $+1$, for the coördinate x_i (or for the coördinate $-x_i$ as the case may be). Hence

(4) *The $\binom{2p+1}{p}$ median points $\xi_{l_1 \ldots l_p}$ may be arranged as the elements of a p-way determinant D' of order $2p+1$ with the lines determined by the ordered subscripts $l_1, \cdots, l_p$. An element in two or more lines of the same index is to be non-existent. With factors of proportionality chosen as above the $\binom{2p+1}{p-1}$ independent linear relations of type* (3) *arise by equating the sum of the elements of any line to zero.*

Let

$$(5) \qquad \left\{ \begin{matrix} i_{p+2} \cdots i_{p+k+1} \\ i_1 \cdots i_{p+1} \end{matrix} \right\}^k \qquad (k = 1, \cdots, p+1)$$

denote the polar of the form $(i_{p+2} t) \cdots (i_{p+k+1} t)$ of order k as to the form $(i_1 t) \cdots (i_{p+1} t)$ of order $p+1$ *without however the usual numerical factor;* and let

$$(6) \qquad \left\{ \begin{matrix} i_{p+2} \cdots i_{p+k+1} \\ i_1 \cdots i_{p+1} \end{matrix} \right\}^k_j$$

denote the value of the polar form (5) for $t = t_j$. We prove the theorem:

(7) *The parametric equation of the norm-curve N^{p-k+1} which lies in the space $\{S\}_k$ of § 2 (14) is, in terms of the parameter t, and for $k = 1, \cdots, p$ as follows:*

$$\Sigma_{i_1 \ldots i_{p+k+1}} \left\{ \begin{matrix} i_1 \cdots i_k \\ i_{k+1} \cdots i_{p+k+1} \end{matrix} \right\}^k \xi_{i_1 \ldots i_k i_{p+k+2} \ldots i_{2p+1}} = 0;$$

where Σ is such as to symmetrize for its subscripts. The equation of the point on M_{2p-1} which is the map of the general ordered $(2p+2)$-ic is, in terms of the equations of the median points,

$$\Sigma_{i_1 \ldots i_{2p+1}} \left\{ \begin{matrix} i_1 \cdots i_p \\ i_{p+1} \cdots i_{2p+1} \end{matrix} \right\}^p_{2p+2} \xi_{i_1 \ldots i_p} = 0.$$

The given equations have coefficients which are polars containing t to an order which ensures the existence of the corresponding norm-curve. In order to prove that N^{p-k+1} is the map of the binary $(2p+2)$-ic with ordered distinct roots $t_{i_1}, \cdots, t_{i_{p+k+1}}$, and coincident roots $t_{i_{p+k+2}} = \cdots = t_{i_{2p+2}} = t_{2p+2} = t$, it is necessary only to show that for $t = t_{i_{p+k+1}}$ in N^{p-k+1} we obtain the point $t = t_{i_{p+k+2}}$ on N^{p-k+2} i. e. that

$$(8) \quad \Sigma_{i_1 \ldots i_{p+k+1}} \left\{ \begin{matrix} i_1 \cdots i_k \\ i_{k+1} \cdots i_{p+k+1} \end{matrix} \right\}^k_{i_{p+k+1}} \xi_{i_1 \ldots i_k i_{p+k+2} \ldots i_{2p+1}}$$
$$= (p-k+3)\, \Sigma_{i_1 \ldots i_{p+k}} \left\{ \begin{matrix} i_1 \cdots i_{k-1} \\ i_k \cdots i_{p+k} \end{matrix} \right\}^{k-1}_{i_{p+k+1}} \xi_{i_1 \ldots i_{k-1} i_{p+k+1} \ldots i_{2p+1}}.$$

For, the norm-curves N in (7) will then be determined by precisely those points which were used in § 2 (14) to define them in succession. In order to prove (8) the following auxiliary formulae are necessary:

$$(9.1) \qquad \left\{ \begin{matrix} i_{p+2} \cdots i_{p+k+1} \\ i_1 \cdots i_{p+1} \end{matrix} \right\}^k_{i_{p+k+1}} = (p-k+2) \left\{ \begin{matrix} i_{p+2} \cdots i_{p+k} \\ i_1 \cdots i_{p+1} \end{matrix} \right\}^{k-1};$$

$$(9.2) \qquad \left\{ \begin{matrix} i_{p+2} \cdots i_{p+k+1} \\ i_1 \cdots i_{p+1} \end{matrix} \right\}^k_{i_1} = -\sum_{r=2}^{r=k+1} \left\{ \begin{matrix} i_{p+2} \cdots i_{p+r-1}\, i_{p+r+1} \cdots i_{p+k+1} \\ i_{p+r}\, i_2 \cdots i_{p+1} \end{matrix} \right\}^{k-1}_{i_1}.$$

These give the value of the polar when t is a root, either of the polarizing form, or of the polarized form. In (9.1) the left member is obtained from the polar form $\left\{ \begin{matrix} i_{p+2} \cdots i_{p+k} \\ i_1 \cdots i_{p+1} \end{matrix} \right\}^{k-1}$ of order $p-k+2$ by operating with $(i_{p+k+1}\, \partial/\partial t)$, and then setting $t = t_{i_{p+k+1}}$. According to Euler's theorem on homogeneous functions, this is the right member. In (9.2) the terms in the polar on the left which contain the factor $(i_1 t)$ vanish for $t = t_{i_1}$. The terms which persist are those which contain i_1 in the form $(i_1 i_{p+r})$ $(r = 2, \cdots, k+1)$. The residual factors are the terms of a polar of next lower order so that the left member can be written as

$$L = \sum_{r=3}^{r=k+1} - (i_{p+r} i_1) \left\{ \begin{matrix} i_{p+2} \cdots i_{p+r-1}\, i_{p+r+1} \cdots i_{p+k+1} \\ i_2 \cdots i_{p+1} \end{matrix} \right\}_{i_1}^{k-1}.$$

On the right the terms of the polars which contain factors of the form $(i_{p+r}\, i_{p+s})$ $(s = 2, \cdots, k+1;\ s \neq r)$ will vanish. For, when r takes the value s, the same terms come in with factor $(i_{p+s}\, i_{p+r})$. The only terms which persist when $t' = i_1$ are those which contain $(i_{p+r} t)$, and these also reduce to L.

Returning to (8), and isolating the index i_{p+k+1}, the sum on the left breaks up into two parts, namely:

$$\Sigma_{i_1 \ldots i_{p+k}} \left\{ \begin{matrix} i_1 \cdots i_{k-1}\, i_{p+k+1} \\ i_k \cdots i_{p+k} \end{matrix} \right\}_{i_{p+k+1}}^{k} \xi_{i_1 \ldots i_{k-1}\, i_{p+k+1} \ldots i_{2p+1}}$$
$$+ \Sigma_{i_1 \ldots i_{p+k}} \left\{ \begin{matrix} i_1 \cdots i_k \\ i_{k+1} \cdots i_{p+k}\, i_{p+k+1} \end{matrix} \right\}_{i_{p+k+1}}^{k} \xi_{i_1 \ldots i_k\, i_{p+k+2} \ldots i_{2p+1}}.$$

On applying (9.1) and (9.2) respectively to these sums, they become

$$(p-k+2)\, \Sigma_{i_1 \ldots i_{p+k}} \left\{ \begin{matrix} i_1 \cdots i_{k-1} \\ i_k \cdots i_{p+k} \end{matrix} \right\}_{i_{p+k+1}}^{k-1} \xi_{i_1 \ldots i_{k-1}\, i_{p+k+1} \ldots i_{2p+1}}$$
$$- \Sigma_{i_1 \ldots i_{p+k}} \sum_{j=1}^{j=k} \left\{ \begin{matrix} i_1 \cdots i_{j-1}\, i_{j+1} \cdots i_k \\ i_j\, i_{k+1} \cdots i_{p+k} \end{matrix} \right\}_{i_{p+k+1}}^{k-1} \xi_{i_1 \ldots i_k\, i_{p+k+2} \ldots i_{2p+1}}.$$

Those negative terms in which the lower line of the brace is $i_k \cdots i_{p+k}$ will have for coefficients

$$- \sum_{r=k}^{r=p+k} \xi_{i_r i_1 \ldots i_{k-1}\, i_{p+k+2} \ldots i_{2p+1}} = \xi_{i_{p+k+1} i_1 \ldots i_{k-1}\, i_{p+k+2} \ldots i_{2p+1}} \quad [\text{cf. } (3)].$$

Thus these negative terms combine with the positive terms to raise the coefficient $(p-k+2)$ to $(p-k+3)$ as on the right of (8).

5. Duality of the invariants (A) with respect to the invariants (B). The final equation of § 4 (7) which expresses the general point of the modular manifold M_{2p-1} in terms of the median points contains all of the median points. The coefficients in it, though formed for isolated t_{2p+2}, are in reality symmetric in all of the roots. The equation might be written in the more symmetric form

$$(1) \qquad \Sigma \left\{ \begin{matrix} i_{p+2} \cdots i_{2p+2} \\ i_1 \cdots i_{p+1} \end{matrix} \right\}^{p+1} \xi_{i_1 \ldots i_{p+1};\, i_{p+2} \ldots i_{2p+2}} = 0.$$

The earlier form has, however, the advantage in that the summation, and the signs of the terms, are more easily defined and we retain it. The coefficients are the apolarity invariants of two complementary factors of order $p+1$ of the binary $(2p+2)$-ic. Each term of such an invariant is itself an invariant (A). In fact

$$(1.1)\qquad \begin{Bmatrix} i_{p+2}\cdots i_{2p+2} \\ i_1\cdots i_{p+1}\end{Bmatrix}^{p+1} = \Sigma\,(i_1 i_{p+2})(i_2 i_{p+3})\cdots(i_{p+1} i_{2p+2})$$

where Σ refers to the $(p+1)!$ permutations of $i_1,\cdots,i_{p+1}$, or of $i_{p+2},\cdots,i_{2p+2}$. We make then a linear transformation from the coördinates x of § 2 (1) to coördinates defined by these apolarity invariants. Again we fix i_{2p+2} to be $2p+2$ and set

$$(2)\qquad x_{i_{p+2}\ldots i_{2p+1}} = \begin{Bmatrix} i_{p+2}\cdots i_{2p+1},\, 2p+2 \\ i_1\cdots i_{p+1}\end{Bmatrix}^{p+1}.$$

With respect to these new coördinates or new invariants (A) we first observe that they satisfy the same set of independent linear relations as the determinant products, or invariants (B) in § 3 (3), or as the median points in § 4 (3), namely:

$$(3)\qquad r_{m_1\ldots m_{p-1}} \equiv x_{k_1 m_1\ldots m_{p-1}} + \cdots + x_{k_{p+2} m_1\ldots m_{p-1}} = 0.$$

For, the identity

$$(4)\qquad \begin{Bmatrix} k_1 \\ k_2\cdots k_{p+2}\end{Bmatrix}^1 + \begin{Bmatrix} k_2 \\ k_1 k_3\cdots k_{p+2}\end{Bmatrix}^1 + \cdots + \begin{Bmatrix} k_{p+2} \\ k_1\cdots k_{p+1}\end{Bmatrix}^1 \equiv 0,$$

is verified by noting that the term $(k_2k_1)(k_3t)\cdots(k_{p+2}t)$ of the first of the polars in (4) is cancelled by the term $(k_1k_2)(k_3t)\cdots(k_{p+2}t)$ of the second, etc. If this identity is polarized further with respect to

$$(m_1t)\cdots(m_{p-1}t)(2p+2,t),$$

the identity (3) is obtained. These relations ensure that the number of linearly independent coördinates x in (2) is not greater than the number, ν, of linearly independent invariants. We ensure that this number is as great as ν by expressing an individual invariant (A) in terms of the invariants x in (2) by the formula,

$$(5)\qquad \frac{(p+2)!}{2}\cdot(12)(34)\cdots(2p+1,2p+2) = \Sigma\,(-1)^p\, x_{135\ldots,2p-1},$$

where the 2^p terms in Σ arise from the one given by the operations of the G_{2^p} generated by the p transpositions, (12), (34), $\cdots$, $(2p-1,2p)$, provided that, when a transposition is applied to a term, the sign of the term is also to be changed. Thus, for $p=2$, and $p=3$, we have

$$12\cdot(12)(34)(56) = x_{13} - x_{23} - x_{14} + x_{24},$$
$$60\cdot(12)(34)(56)(78) = -x_{135} + x_{235} + x_{145} - x_{245} + x_{136} - x_{236} - x_{146} + x_{2\cdot}$$

To prove (5) we note first that the transposition (12) changes the sign

on the right, whence Σ contains the factor (12); and, by the same argument, Σ contains the factors $(34), \cdots, (2p-1, 2p)$. As a function of the differences of the roots linear in each root Σ must then also contain the factor $(2p+1, 2p+2)$. In order to verify however that Σ does not vanish identically, and at the same time to check the numerical factor in (5), let the odd-numbered roots all be equal to r and the even numbered roots all be equal to s. Then the left member of (5) has the value $(rs)^p \cdot (p+2)!/2$. Each term on the right has the form

$$(-1)^k \begin{Bmatrix} r\,r \cdots r\; s\,s \cdots s \\ s\,s \cdots s\; r\,r \cdots r \end{Bmatrix}^{p+1}$$

where k is the number of r's in the upper line and $k = 0, 1, \cdots, p$ since $t_{2p+2} = s$ always. This value will however arise from $\binom{p}{k}$ terms since in general the $(p-k)$ roots s in the upper line other than the last may be any of $t_2, t_4, \cdots, t_{2p}$. In forming the apolarity invariant the roots s, r in the lower line are distributed in any order with respect to the roots r, s in the upper line so that the non-vanishing terms contribute

$$k!\,(sr)^k \cdot (p+1-k)!\,(rs)^{p+1-k} = (-1)^k\, k!\,(p+1-k)!\,(rs)^{p+1}.$$

The value of Σ is therefore $\left[\sum_{k=0}^{k=p} \binom{p}{k}\, k!\,(p+1-k)!\right](rs)^{p+1}$. Since $\binom{p}{k} = p!/k!\,(p-k)!$ this can be written as

$$p!\,(rs)^{p+1}\left[\sum_{k=0}^{k=p}(p+1-k)\right] = p!\,(rs)^{p+1}\cdot(p+1)(p+2)/2$$

which coincides with the left member.

For values of p beyond three no convenient choice of ν independent variables $x_{m_1 \ldots m_p}$, nor of independent combinations of them, seems to exist. Even for $p = 2, 3$ the best procedure is to select a set of $\nu + 1$ combinations whose sum is identically zero (for $p = 2$ cf. [1], I, p. 167; for $p = 3$, cf. Part II of this account). Thus, in the general case, it seems best to retain all the variable point coördinates $x_{m_1 \ldots m_p}$, and therefore to carry along the linear relations (3) on these coördinates. The question then arises as to the nature of the dual coördinates ξ in $S_{\nu-1}$. If the variables were independent we should have dual coördinates x, ξ with the incidence condition $\Sigma\, x_i \xi_i = 0$. When the linear relations are introduced, the coefficients ξ of this condition can be modified into

$$\Sigma_i x_i \xi_i + \Sigma \lambda_{m_1 \ldots m_{p-1}} r_{m_1 \ldots m_{p-1}} = 0.$$

If we write this as $\Sigma\, x_i \xi'_i$, the values of the ξ'_i are

(6) $$\xi'_{m_1 \ldots m_p} = \xi_{m_1 \ldots m_p} + \lambda_{m_2 \ldots m_p} + \lambda_{m_1 m_3 \ldots m_p} + \cdots + \lambda_{m_1 m_2 \ldots m_{p-1}},$$

where the constants λ are at our disposal. It is natural to attempt to choose these $\binom{2p+1}{p-1}$ constants λ so that the coördinates ξ' satisfy the same system of $\binom{2p+1}{p-1}$ linear relations as the coördinates x in (3). If

$$(7) \qquad r_{m_2 \ldots m_p} \equiv \xi'_{k_1 m_2 \ldots m_p} + \xi'_{k_2 m_2 \ldots m_p} + \cdots + \xi'_{k_{p+2} m_2 \ldots m_p} = 0,$$

then from (6) there follows that

$$\begin{aligned}(8) \quad (p+2)\lambda_{m_2 \ldots m_p} &+ \lambda_{k_1 m_3 \ldots m_p} + \lambda_{k_1 m_2 m_4 \ldots m_p} + \cdots + \lambda_{k_1 m_2 m_3 \ldots m_{p-1}} \\ &+ \lambda_{k_2 m_3 \ldots m_p} + \lambda_{k_2 m_2 m_4 \ldots m_p} + \cdots + \lambda_{k_2 m_2 m_3 \ldots m_{p-1}} \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ &+ \lambda_{k_{p+2} m_3 \ldots m_p} + \lambda_{k_{p+2} m_2 m_4 \ldots m_p} + \cdots + \lambda_{k_{p+2} m_2 m_3 \ldots m_{p-1}} \\ = - \xi_{k_1 m_2 \ldots m_p} &- \xi_{k_2 m_2 \ldots m_p} - \cdots - \xi_{k_{p+2} m_2 \ldots m_p}.\end{aligned}$$

If the $\binom{2p+1}{p-1}$ equations (8) in the $\binom{2p+1}{p-1}$ constants λ can be solved for the λ's in terms of the ξ's, the relations (7) can be satisfied. That the solution is possible (i. e., that the determinant of the system is not zero) can be proved by showing that any individual $\lambda_{m_2 \ldots m_p}$ can be obtained by taking proper linear combinations of the left members of equations (8). For, if that equation derived from (8) by the interchange of m_2 and k_1 be written, the difference of the two left members is

$$\begin{aligned}(9) \quad (\lambda_{m_2} - \lambda_{k_1})\{(p+1)\lambda_{m_3 \ldots m_p} &+ \lambda_{k_2 m_4 \ldots m_p} + \cdots + \lambda_{k_2 m_3 \ldots m_{p-1}} \\ &+ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ &+ \lambda_{k_{p+2} m_4 \ldots m_p} + \cdots + \lambda_{k_{p+2} m_3 \ldots m_{p-1}}\}.\end{aligned}$$

In (9) the multiplication is symbolic, i. e., $\lambda_{m_2}\lambda_{m_3} \cdots \lambda_{m_p} = \lambda_{m_2 \ldots m_p}$. The part of (9) within the brace is the left member of (8) formed for $p' = p - 1$. Assuming that this latter system can be solved for $\lambda_{m_3 \ldots m_p}$, the original system can be solved for $\lambda_{m_2 m_3 \ldots m_p} - \lambda_{k_1 m_3 \ldots m_p}$, i. e., for the difference of $\lambda_{m_2 \ldots m_p}$ and any λ which has $p-2$ subscripts in common with it. On adding such properly chosen differences to (8) the resulting left member is a multiple of $\lambda_{m_2 \ldots m_p}$. Thus the system (8) can be solved for the λ's if it can be solved in the earliest case $p = 2$. For this case the five left members have the typical form,

$$4\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5$$

and the possibility of the solution is obvious.

At the moment we do not need the explicit values of the new coördinates ξ' in (6); but they are useful later, and therefore we give the solution of the system (8) and the resulting values of the ξ''s in (6). So long as only linear expressions are in question, it is legitimate to use the symbolic multiplication,

$$(10)\qquad \begin{aligned}\xi_{123\ldots} &= \xi_1\xi_2\xi_3\cdots = \xi_1\xi_{23}\cdots = \cdots;\\ \lambda_{123\ldots} &= \lambda_1\lambda_2\lambda_3\cdots = \lambda_1\lambda_{23}\cdots = \cdots.\end{aligned}$$

If the $2p+1$ subscripts are divided into complementary sets $m_1,\cdots,m_{p-1}$ and $k_1,\cdots,k_{p+2}$, and if Σ, S refer respectively to sums which symmetrize with respect to these sets, the equation (8) can be written in the form

$$(11)\qquad \frac{(p+2)!}{2}\,[(p+2)\lambda_{m_1\ldots m_{p-1}} + \Sigma\lambda_{m_1\ldots m_{p-2}}S\lambda_{k_1}] = -\frac{(p+2)!}{2}\,\xi_{m_1\ldots m_{p-1}}S\xi_{k_1}.$$

Due to the symmetry of the system (8) in all the indices we may assume that the solution whose existence was proved above has the form,

$$(12)\qquad \frac{(p+2)!}{2}\,\lambda_{m_1\ldots m_{p-1}} = \sum_{j=0}^{p-1} a_j\Sigma\,\xi_{m_1\ldots m_{p-1-j}}S\xi_{k_1\ldots k_{j+1}}.$$

If we substitute this value in (11), and note that the coefficient of $\xi_{m_1\ldots m_{p-1-j}}\xi_{k_1\ldots k_{j+1}}$ on the right is zero, except that for $j=0$ it is $-(p+2)!/2$, then $a_0,\cdots,a_{p-1}$ are determined by

$$(13)\qquad \begin{aligned}&(2p+1)a_0 + (p-1)(p+1)a_1 = -(p+2)/2,\\ &j(j+1)a_{j-1} + (j+1)(2p-2j+1)a_j\\ &\qquad\qquad + (p-j-1)(p-j+1)a_{j+1} = 0\\ &[a_p = 0,\ j = 1,\cdots,p-1].\end{aligned}$$

An easy verification shows that

$$a_0 = -(p+1)!/2,\quad a_j = (-1)^{j-1}j!\,(p-j+1)!/2$$

or

$$(14)\qquad \frac{(p+2)!}{2}\,\lambda_{m_1\ldots m_{p-1}} = -\frac{(p+1)!}{2}\,\xi_{m_1\ldots m_{p-1}}S\xi_{k_1} + \sum_{j=1}^{p-1}\frac{(-1)^{j-1}j!\,(p-j+1)!}{2}\,\Sigma\xi_{m_1\ldots m_{p-1-j}}S\xi_{k_1\ldots k_{j+1}}.$$

Let the equation (6) be written in the form,

$$(15)\qquad \xi'_{m_1\ldots m_p} = \xi_{m_1\ldots m_p} + \mu_{m_1\ldots m_p}.$$

Then

$$(16)\qquad \frac{(p+2)!}{2}\,\mu_{m_1\ldots m_p} = \sum_{j=0}^{p}\alpha_j\;\Sigma'\xi_{m_1\ldots m_{p-j}}S'\xi_{k_1\ldots k_j},$$

where Σ', S' refer to complementary sets $m_1,\cdots,m_p$; $k_1,\cdots,k_{p+1}$. On substituting the values of λ as given in (14) in the value of μ obtained from (6), we find that

$$\alpha_j = ja_{j-1} + (p-j)a_j \qquad (j = 0,\cdots,p).$$

Using the values of the a_j in (14) this yields

$$\alpha_0 = -\frac{(p+1)!\,p}{2}, \quad \alpha_j = (-1)^j(p-j+1)!\,j! \qquad (j=1,\cdots,p),$$

whence

$$(17) \qquad \frac{(p+2)!}{2}\,\xi'_{m_1\ldots m_p} = (p+1)!\,\xi_{m_1\ldots m_p} + \sum_{j=1}^{p}(-1)^j j!\,(p-j+1)!\,\Sigma'\xi_{m_1\ldots m_{p-j}}\,S'\xi_{k_1\ldots k_j}.$$

We prove now that

(18) *With* $x_{m_1\ldots m_p} = \left\{ \begin{matrix} m_1\cdots m_p, 2p+2 \\ k_1\cdots k_{p+1} \end{matrix} \right\}^{p+1}$ [cf. (2)], *and*

$$\xi_{m_1\ldots m_p} = \epsilon(k_1\cdots k_{p+1})(m_1\cdots m_p, 2p+2) \quad [\text{cf. } \S\, 3\ (2)],$$

the aggregate of projectively distinct and ordered binary $(2p+2)$-ics is mapped upon the points x of a manifold $M_{2p-1}(x)$ in $S_{\nu-1}$, and also upon the spaces ξ of a manifold $M_{2p-1}(\xi)$ in $S_{\nu-1}$ which is dually related to $M_{2p-1}(x)$ in such wise that the SAME *ordered $(2p+2)$-ic determines* INCIDENT *point x and space ξ. Under permutation of the roots of the $(2p+2)$-ic the point x and space ξ are subject to the operations of the* SAME *collineation group, $C_{(2p+2)!}$, for which also the manifolds M are invariant.*

For, the variables ξ, defined as in (18), satisfy according to § 3 (3) the same set of linear relations as the variables x [cf. (3)]; and therefore, as proved above, they may be taken as dual coördinates with the incidence condition

$$(19) \quad \Sigma x_{m_1\ldots m_p}\xi_{m_1\ldots m_p} = 0 \quad (m_1,\cdots,m_p = 1,\cdots,2p+1;\ m_i \neq m_j).$$

The x's are of degree $(p+1)$ in the differences of the roots, the ξ's of degree $p(p+1)$, so that $x_i\xi_i$ is of degree $(p+1)^2$. On interchange of t_{m_i}, t_{m_j} the ξ's are permuted with a change of sign throughout, while the x's are merely permuted. Hence $(x\xi)$ must contain (m_im_j) as a factor, and thus must contain as a factor the $p(2p+1)$ differences which is impossible since its degree is $(p+1)^2$. Thus $(x\xi)$, as a function of the roots, must vanish. When the roots are permuted, the variables x, ξ are permuted in the same way, and the two permutations are dual forms of the same collineation in $S_{\nu-1}$.

When the $M_{2p-1}(x)$ is regarded as the map of the linear space $S_{2p-1}(y)$ as in § 2 (7), the ordered binary $(2p+2)$-ic determines a point y in S_{2p-1} and also the map x of this point y on $M_{2p-1}(x)$ in $S_{\nu-1}$; but it determines also as in (18) a linear space ξ in $S_{\nu-1}$, and therefore a particular member of the linear mapping system in $S_{2p-1}(y)$. We determine this particular

member in the following way and obtain thereby a new interpretation of the incidence relation.

With basal coördinates y in $S_{2p-1}(y)$ such that

$$y_1 + y_2 + \cdots + y_{2p+1} = 0, \tag{20}$$

and $2p+1$ basis points of which the i-th is

$$y_i = -2p, \qquad y_j = 1 \qquad (j \neq i), \tag{21}$$

le equation in terms of the parameter t of the norm-curve N^{2p-1}, which asses through the base points with parameters $t_1, \cdots, t_{2p+1}$ respectively, is

$$y_i - y_j = (ij)\Pi/(it)(jt), \qquad \Pi = (1t)(2t)\cdots(2p+1, t); \tag{22}$$

$$2py_i = \Sigma_j (ij)\Pi/(it)(jt).$$

his is in effect the space $\{S\}_1 = S_{2p-1}(y)$ of § 4 (7) when the root t_{2p+2} is transformed to ∞ and the other roots are transformed into $y_1, \cdots, y_{2p+1}$.

For dual coördinates in $S_{2p-1}(y)$ we use $\eta_1, \cdots, \eta_{2p+1}$ where

$$\eta_1 + \eta_2 + \cdots + \eta_{2p+1} = 0; \tag{23}$$

and

$$\eta_1 y_1 + \eta_2 y_2 + \cdots + \eta_{2p+1} y_{2p+1} = 0. \tag{24}$$

the incidence condition of point y and space η. The norm curve N^{2p-1} is ocus of spaces η as well as a locus of points y. The dual parametric equation N^{2p-1} is

$$\eta_i = (-1)^{i+1}(123\cdots, i-1, i+1, \cdots, 2p+1)(it)^{2p-1} \qquad (i = 1, \cdots, 2p-1). \tag{25}$$

ve observe first of all that the linear relation (23) is satisfied due to tl ear relation,

$$\sum_{i=1}^{2p+1} (-1)^{i+1}(12\cdots, i-1, i+1, \cdots, 2p+1)(it)^{2p-1} = 0, \tag{26}$$

among $2p+1$ perfect $(2p-1)$-th powers of linear forms. Secondly we prove that the space $\eta(t)$ in (25) will cut the curve $y(t')$ in (22) in $2p-1$ points determined by $(t't)^{2p-1} = 0$. Indeed the incidence condition (24) can be modified, by using (20) and then (22), to read

$$\sum_{i=1}^{2p} \eta_i [y_i(t') - y_{2p+1}(t')]$$

$$= \sum_{i=1}^{2p} (-1)^{i+1}(12\cdots, i-1, i+1, \cdots, 2p+1)(it)^{2p-1} \times (i, 2p+1)\Pi(t')/(it')(2p+1, t')$$

$$= \prod_{i=1}^{2p} (i, 2p+1) \cdot \sum_{i=1}^{2p} (-1)^{i+1}(12\cdots, i-1, i+1, \cdots, 2p, t') \cdot (it)^{2p-1} = 0.$$

This reduces, by using the relation (26) in which t_{2p+1} is replaced by t', to $-[\prod_{i=1}^{2p}(i, 2p+1)]\cdot(123\cdots 2p)\cdot(t't)^{2p-1}=0$. Thus the equation $(\eta y)=0$ of the space $\eta(t)$ of the N^{2p-1} in (25) is

$$(27)\quad y_1(234\cdots, 2p+1)(1t)^{2p-1}-y_2(134\cdots, 2p+1)(2t)^{2p-1} \\ +\cdots+y_{2p+1}(123\cdots, 2p)(2p+1, t)^{2p-1}=0.$$

For given t, (27) is the equation in variables y of the space η which has $(2p-1)$-point contact with N^{2p-1} at the point t; for given y, (27) is the equation in variable t which determines the $2p-1$ parameters t of spaces of N^{2p-1} on the point y.

A binary form of odd order $2p-1$ such as (27) has a covariant of degree p, order p, and weight $p(p-1)$, namely, its *canonizant*, such that the form can be expressed as a sum of the $(2p-1)$-th powers of the linear factors of the canonizant. If the form can be expressed as a sum of less than p such powers the canonizant vanishes identically. If then the form is

$$k_1(1t)^{2p-1}+k_2(2t)^{2p-1}+\cdots+k_p(pt)^{2p-1},$$

the canonizant is, to within a numerical factor,

$$k_1k_2\cdots k_p(12\cdots p)^2(1t)(2t)\cdots(pt).$$

Hence the canonizant of (27) is made up of a sum of such products, each product arising from p terms of (27). We examine the above typical product when $k_1=(23\cdots, 2p+1)y_1$, etc., and observe first that it contains $[(\Delta_{2p+1})^{1/2}]^{p-1}$ as a factor where $(\Delta_{2p+1})^{1/2}$ is the product of differences, $(12\cdots, 2p+1)$. The residual factor is

$$y_1y_2\cdots y_p(12\cdots pt)(p+1,\cdots, 2p+1)=d_{12\ldots p}\,y_1y_2\cdots y_p.$$

There is in addition a sign depending on the number of minus signs occurring in the p terms in (27). This sign will change as the p terms selected from (27) change, but the change is precisely that accounted for in the definition of $d_{12\ldots p}$ in § 3 (2). Hence

(28) *The canonizant of the binary $(2p-1)$-ic in (27) is, to within a numerical factor and the factor,* $[(\Delta_{2p+1})^{1/2}]^{p-1}$,

$$\Sigma_{12\ldots p}\, d_{12\ldots p}\, y_1y_2\cdots y_p.$$

This expression for the canonizant can be modified so as to take the form (19) of the incidence condition in $S_{\nu-1}$. For, the $d_{12\ldots p}$ satisfy according to § 3 (3) the same system of linear relations as the coördinates

$x_{12\ldots p}$ in (3), and therefore their coefficients may be modified as in (17) the $\xi_{12\ldots p}$ were modified into $\xi'_{12\ldots p}$. Thus the canonizant (28) can be written as

$$\frac{(p+2)!}{2}\Sigma_{12\ldots p}\, d_{12\ldots p}\, y_1 y_2\cdots y_p$$
$$=\Sigma_{12\ldots p}\, d_{12\ldots p}\{(p+1)!\, y_1y_2\cdots y_p$$
$$+\sum_{j=1}^{p}(-1)^j j!\,(p-j+1)!\,\Sigma'' y_1\cdots y_{p-j}\, S'' y_{p+1}\cdots y_{p+j}\},$$

where Σ'' refers to a symmetrization as to $1,\cdots,p$, and S'' as to $p+1,\cdots,2p+1$. The coefficient of $d_{12\ldots p}$ is now

$$\text{(29)}\qquad \begin{Bmatrix} y_1\cdots y_p\ \infty \\ y_{p+1}\cdots y_{2p+1}\end{Bmatrix}^{p+1}.$$

For, the value of this polar is

$$S''(y_1-y_{p+1})(y_2-y_{p+2})\cdots(y_p-y_{2p})$$

i. e. it is a sum of $(p+1)!$ products. In each product there is a term $y_1y_2\cdots y_p$. Moreover in the sum there are $(p+1)!\binom{p}{j}$ terms of the type $(-1)^j y_1\cdots y_{p-j}\, y_{p+1}\cdots y_{p+j}$, and $\Sigma'' y_1\cdots y_{p-j}\, S'' y_{p+1}\cdots y_{p+j}$ contains $\binom{p}{j}\binom{p+1}{j}$ terms of this type, whence each term arises $(p-j+1)!\,j!$ times. Hence

(30) *To within an additional numerical factor the canonizant* (28) *can be written as*

$$\Sigma_{12\ldots p}\, d_{12\ldots p}\begin{Bmatrix} y_1\cdots y_p\ \infty \\ y_{p+1}\cdots y_{2p+1}\end{Bmatrix}.$$

We recall that the $d_{12\ldots p}$ in (30) have been formed for $t_1,\cdots,t_{2p+1}$ and $t_{2p+2}=t$. Their polar coefficients have been formed from $y_1,\cdots,y_{2p+1},\infty$ which are projective to $t_1,\cdots,t_{2p+1},t$. Hence the canonizant is the form taken by the incidence condition (19) [cf. also (18)]. This leads to a determination of the nature of the p-ic spread (30), and of the place of the linear space ξ in the mapping of $S_{2p-1}(y)$ on $M_{2p-1}(x)$ in $S_{\nu-1}$. When $t_1,\cdots,t_{2p+1}$ are given, the norm-curve N^{2p-1} in $S_{2p-1}(y)$ is determined, and for given y in (27) the parameters t of the $2p-1$ spaces of N^{2p-1} on y are determined. The canonizant of this $(2p-1)$-ic furnishes the p values $t=\tau_1,\cdots,\tau_p$ such that y is on an S_{p-1} which cuts N^{2p-1} at the points $t=\tau_1,\cdots,\tau_p$, say *a p-secant* S_{p-1} of N^{2p-1}. If then this canonizant is set equal to zero, for given t we obtain the locus in variables y of the ∞^{p-1} p-secant S_{p-1}'s of N^{2p-1} which are on the point t of N^{2p-1}. Hence

(31) *In* (18) *the ordered* $(2p+2)$*-ic determines an incident point* x *and space* ξ *in* $S_{\nu-1}$. *When the* $(2p+2)$*-ic is first mapped on a point* y *in* $S_{2p-1}(y)$, *and* y *is then mapped upon* x *in* $S_{\nu-1}$ [cf. § 2 (7)], *the linear space* ξ *arises from a particular member of the mapping system defined at* y *by the basic norm-curve* N^{2p-1} *on* y. *This member is a cone of order* p, *the locus of the* ∞^{p-1} p*-secant spaces* S_{p-1} *of* N^{2p-1}, *each of which is on* y. *A section of this cone is the locus in* S_{2p-2} *of the* ∞^{p-1} $(p-1)$*-secant spaces* S_{p-2} *of a norm-curve* N^{2p-2}.

6. Applications to W_3 and W_p. We close this initial article with a contrast of the incidence condition described in § 5 (18), (31) for the two cases $p=2$ and $p=3$. This condition is bilinear in the invariants (A) and (B). The invariants (B) however can be expressed as polynomials of degree p in the invariants (A) [cf. [6]]. When the invariants (A) are expressed as polynomials of degree p in the coördinates of a point y in $S_{2p-1}(y)$, and the invariants (B) in terms of a point y' in $S_{2p-1}(y)$, this incidence condition becomes

$$f(y^p, y'^{p^2})=0, \tag{32}$$

where f is a double form of the degrees indicated. According to (31) this equation expresses that, for given y', y is a point on a p-secant space S_{p-1} through y' of the norm-curve N^{2p-1} determined by the set of $2p+2$ points $P'^{\,2p-1}_{2p+2}$, made up of the base of the mapping system (say $p_1, \cdots, p_{2p+1}$) and y'. In the notation introduced in connection with the generalized Weddle W_p ([4], p. 449), and for $y'=p_{2p+2}$, $f=0$ is the equation of the P-locus, $\pi_{1,2,\ldots,2p+1}$.

For $p=2$, the incidence condition, $f(y^2, y'^4)=0$, is for given y' a quadric cone with generators on y' and the points of the N^3 on $P'^{\,3}_6=p_1, \cdots, p_5, y'$. Thus it is the condition that $P_6^3=p_1, \cdots, p_5, y$, be on a quadric cone with vertex at y'. Hence for given y it is the equation in variables y' of the Weddle surface determined by P_6^3. The symmetry of the condition in the six points of P_6^3 is notable.

For $p=3$, this symmetry does not persist. The $f(y^3, y'^9)=0$ is, for given y', the locus of trisecant planes on y' of the N^5 on $P'^{\,5}_8=p_1, \cdots, p_7, y'$. Therefore it is the condition that $P_8^5=p_1, \cdots, p_7, y$ be projected from y' into $Q_8^4=q_1, \cdots, q_7, z$ where z is a point on the bisecant locus B of the N^4 determined by $q_1, \cdots, q_7$. We have obtained [[4], p. 478 (67e)] the equation of B in terms of $q_1, \cdots, q_7$ in the form,

$$\begin{aligned}\Sigma(14567)(23567)(24157)(24367)&\\ \times(13257)(13467)(2413z)(2456z)(1356z)&=0,\end{aligned} \tag{33}$$

where Σ refers to the six terms obtained by permutation of 1, 2, 6 (or 3, 4, 5). By the Clebsch principle of transference,

$$(34)\qquad f(y^3, y'^9) = \Sigma(14567y')(23567y')(24157y')(24367y') \times (13257y')(13467y')(2413yy')(2456yy')(1356yy') = 0.$$

Thus the condition $f(y^3, y'^9) = 0$ is of degree six in $p_1, \cdots, p_7$ and of degree three in y.

For $p = 2$ and $P_6^3 = p_1, \cdots, p_5, p_6 = y$, the surface $f(y^2, y'^4) = 0$ is precisely the Weddle W_2. For $p = 3$ and $P_8^5 = p_1, \cdots, p_7, p_8 = y$, we ask whether the spread $f(y^3, y'^9) = 0$ contains the Weddle 3-way, W_3; and, if not, how it cuts W_3. If we follow out the argument ([4], pp. 478-9) from which the equation (33) was obtained, in the form in which it transfers, as in (34), to S_5, we obtain

$$(35)\qquad H_{24}H_{13}G_{148}G_{238} - H_{23}H_{14}G_{138}G_{248} = G_{567}\cdot f(y^3, y'^9).$$

In this G_{123} is the S_4 on $p_4, \cdots, p_8$, and H_{78} is the 4-way Weddle cone of order 4 with 4-fold line p_7p_8 obtained by the Clebsch principle from the W_2 on points $q_1, \cdots, q_6$ in S_3. With respect to the indices $1, \cdots, 6; 7, 8$ the F-loci of the third kind of P_8^5 [cf. [4], § 1] divide into types $\pi^{(3)}_{178}$, $\pi^{(3)}_{128}$, $\pi_8^{(3)}$, $\pi_1^{(3)}$ where $\pi^{(3)}_{ijk}$ is the plane $p_ip_jp_k$, and $\pi_i^{(3)}$ is the quartic 2-way cone of bisecants of N^5 on p_i. Then H_{78} contains $\pi^{(3)}_{178}$ doubly, $\pi^{(3)}_{128}$ and $\pi_8^{(3)}$ simply, and does not contain $\pi_1^{(3)}$ [cf. [4] (78f)]. Also H_{78} contains the surface $V_2^{(5)}$ [cf. [4] (76)] with respect to which we wish to correct and amplify the theorem [[4] (82)] to read as follows:

(36) *W_3 contains the surface $V_2^{(5)}$, a 2-way of order 26 with 4-fold points at P_8^5 and with simple curve N^5 and simple lines p_ip_j.*

The correction is with respect to the order 26 which was given as 15. In fact $V_2^{(5)}$ is projected from p_8 into $V_2^{(4)}$ of order 11 doubly covered as will appear later in connection with (39). This double covering of $V_2^{(4)}$ was overlooked in ([4]).

In order to find the intersection of (35) with W_3 we examine first the intersection of H_{78} with W_3. Since W_3 has the order 19 with 9-fold points at P_8^5, triple curves N^5 and p_ip_j, and simple F-loci of the third kind of P_8^5 [[4] (81)], H_{78} cuts W_3 in a 2-way of order 76 from which the six $\pi^{(3)}_{178}$ separate doubly, the thirty $\pi^{(3)}_{128}$ and the two cones $\pi_7^{(3)}$, $\pi_8^{(3)}$ separate simply, the residual intersection is precisely $V_2^{(5)}$ of order 26 with 4-fold points at P_8^5 and simple curve N^5. Since $V_2^{(5)}$ is itself invariant under the Cremona group $G_{2.2^6}$ of W_3, it must also contain the lines p_ip_j conjugate to N^5 simply.

It is now clear from the form of the left member of (35) that $f(y^3, y'^9) = f(p_8^3, y'^9) = 0$ contains $V_2^{(5)}$ doubly, the 35 planes $\pi^{(8)}_{123}$ doubly, and the 21 planes $\pi^{(3)}_{128}$ as well as the seven cones $\pi_1^{(3)}$ simply. Since $2.26 + 35.2 + 21 + 7.4 = 9.19$, the total intersection of $f(p_8^3, y'^9) = 0$ with W_8 is accounted for. Moreover the left member of (35) does not contain W_8. For, according to [4 (78h)], the following octavic spreads contain W_8:

(37) $\quad \pm H_{23}H_{14}/(23)(14) = \pm H_{31}H_{24}/(31)(24) = \pm H_{12}H_{34}/(12)(34).$

Hence on W_8 this left member is equivalent to

$$(31)(24)G_{148}G_{238} \pm (23)(14)G_{138}G_{248} = G_{12348},$$

the quadric with nodes at p_5, p_6, p_7 and simple points at $p_1, \cdots, p_4, p_8$ [cf. [4] (78e)]. This quadric however does not contain W_8 but rather meets W_8 in certain F-loci of the third kind and in a 9-ic 2-way which is on G_{56} also [4, p. 489]. Hence

(38) *The 9-ic locus, $f(p_8^3, y'^9) = 0$, which is determined by the incidence condition of* § 5 (18), (31) *has a nodal locus $V_2^{(5)}$. It also contains all the F-loci of P_8^5 of the third kind except the quartic cone $\pi_8^{(3)}$, and in particular contains the 35 planes of type $\pi^{(3)}_{123}$ doubly.*

Thus the property of W_2 with respect to this incidence condition extends not to W_8 but rather to the surface $V_2^{(5)}$ on W_8. This is a situation which will frequently recur.

We have attached the ∞^p points of W_p to the ∞^p projectively distinct but birationally equivalent curves H_p^{p+2} of genus p and order $p+2$ with p-fold point at O and branch points $r_1, \cdots, r_{2p+2}$ for which the pencil of lines from O to the branch points is projectively given. We inquire therefore as to the projective peculiarity of H_p^{p+2} when the point of W_p is on $V_2^{(2p-1)}$.

In S_{2p-1} a normal elliptic curve E^{2p} depends upon $4p^2$ constants. Since it is $2p-2$ conditions that E^{2p} pass through a given point, there are ∞^4 curves E^{2p} on P^{2p-1}_{2p+2}. If such a curve E^{2p} has two nodes it breaks up into two rational norm-curves bisecant to each other. These degenerate E^{2p}'s are distributed in a variety of families, each ∞^2. A particular set of $2p+2$ of the families is of special interest. One such family consists of ∞^2 curves each made up of a line on p_i and an N^{2p-1} bisecant to the line and on the remaining points of P^{2p-1}_{2p+2}. As i runs from 1 to $2p+2$ the $2p+2$ families are obtained. Each of these $2p+2$ families has the property that it is invariant under the Cremona group $G_{2.2^{2p}}$ of W_p. This indeed is almost evident for the generators I_{ij} of $G_{2.2^{2p}}$. If n_1, n_2 are the nodes of an E^{2p} in

the family determined by p_i, then from n_1 (or n_2) the set P^{2p-1}_{2p+2} is projected into a set Q^{2p-1}_{2p+2} in S_{2p-2} which lies on an N^{2p-2}. But this is the condition that n_1 be a point on $V_2^{(2p-1)}$ [cf. [4] (76)].

The existence of this family of degenerate E^{2p}'s explains the double covering of the projection of $V_2^{(2p-1)}$ upon $V_2^{(2p-2)}$ from the point p_i. By beginning with a point of $V_2^{(2p-1)}$ and projecting it, and the new points so obtained, from the points of P^{2p-1}_{2p+2} upon $V_2^{(2p-1)}$ a closed system of 2.2^{2p} points is obtained on $V_2^{(2p-1)}$ analogous to Baker's closed set of 2.16 points on W_2. The closure is a consequence of the invariance of each of the $2p+2$ families under the group of W_p.

(39) *A line through the $2(p-1)$-fold point p_i of $V_2^{(2p-1)}$ to a point of $V_2^{(2p-1)}$ meets $V_2^{(2p-1)}$ in another point. By repeated projections from points p_i a point y on $V_2^{(2p-1)}$ gives rise to a closed set of 2.2^{2p} points on $V_2^{(2p-1)}$ which divide into two sets of 2^{2p} points, each a conjugate set under the group of the W_p.*

Here we again have a property of W_2 which generalizes into a property of $V_2^{(2p-1)}$ rather than a property of W_p.

The essential condition on a point y of $V_2^{(2p-1)}$ is that the set P^{2p-1}_{2p+2} is projected from y upon a set Q^{2p-2}_{2p+2} on an N^{2p-2}. But P^{2p-1}_{2p+2}, y are associated with the planar set of points $r_1, \cdots, r_{2p+3}, O$. Hence the projection, Q^{2p-2}_{2p+2}, of P^{2p-1}_{2p+2} from y is associated with the planar set $r_1, \cdots, r_{2p+2}$ [1, I, p. 158 (6)]. But Q^{2p-2}_{2p+2} is a set on a rational norm-curve N^{2p-2} whence $r_1, \cdots, r_{2p+2}$ is also a set on a rational norm-curve, i. e., a conic, with parameters on the conic projective to the parameters of Q^{2p-2}_{2p+2} on N^{2p-2} [7, p. 12, Theo. (10)]. If $H_p^{p+2} \equiv x_0^2 f_p + 2x_0 f_{p+1} + f_{p+2} = 0$, the branch points of H_p^{p+2} are on the conic $H_0^2 \equiv x_0^2 g_0 + 2x_0 g_1 + g_2 = 0$ if

$$(40) \qquad g_0 f_{p+2} - 2g_1 f_{p+1} + g_2 f_p \equiv 0.$$

This identical vanishing of the $(p+2)$-ic (40) imposes $p+3$ conditions, of which only five can be satisfied by proper choice of g_0, g_1, g_2, leaving $2p-2$ projective conditions to be satisfied by f_p, f_{p+1}, f_{p+2}. These conditions confine the representative point y on W_p to the locus $V_2^{(2p-1)}$. Hence

(41) *The locus $V_2^{(2p-1)}$ on W_p is that locus whose points correspond to hyperelliptic curves H_p^{p+2} whose $2p+2$ branch points are on a conic.*

It may be observed that, for $p=3$, the 9 points $r_1, \cdots, r_8, O$ are the base points of a pencil of cubics. Thus O is determined in general by

$r_1, \cdots, r_8$. If however $r_1, \cdots, r_8$ are on a conic, O may be any point of the plane. Let $s_1, \cdots, s_8$ be the parameters of $r_1, \cdots, r_8$ on their conic. In S_5 let $s_1, \cdots, s_7, s_8$ be the parameters of the points $p_1, \cdots, p_7, n_2$ on the N'^5 through them. On the quartic 2-way cone of lines from n_2 to the points of N'^5 let p_8 be any point, and let the line p_8n_2 cut N'^5 again at n_1. Then the Weddle W_3 determined by $p_1, \cdots, p_8$ will contain a $V_2^{(5)}$ on n_1, n_2. On the N^5 determined by $p_1, \cdots, p_8$ these points will have parameters $t_1, \cdots, t_8$. Then there will be a unique point O in the plane for which the octavic pencil of lines from O to $r_1, \cdots, r_8$ will be projective to $t_1, \cdots, t_8$. As p_8 runs over the above quartic cone, the point O runs over the plane.

On W_p there will be similar manifolds of higher dimension, $V_4, V_6, \cdots$, which map curves H_p^{p+2} whose branch points are respectively on a cubic with O, on a quartic with node at $O, \cdots$. Doubtless these further manifolds in the higher cases will have properties as interesting as those of $V_2^{(2p-1)}$.

REFERENCES

[1] A. B. Coble, "Point Sets and Allied Cremona Groups," *Transactions of the American Mathematical Society*, I, Vol. 16 (1915), pp. 155-198; II, Vol. 17 (1916), pp. 345-385; III, Vol. 18 (1917), pp. 331-372.

[2] A. B. Coble, "Finite Geometry . . . Characteristic Theory of the Odd and Even Theta Functions," *Transactions of the American Mathematical Society*, Vol. 14 (1913), pp. 241-276.

[3] A. B. Coble, "Algebraic Geometry and Theta Functions," *Colloquium Lectures of the American Mathematical Society*, Vol. 10, New York (1929).

[4] A. B. Coble, "A Generalization of the Weddle Surface . . .," *American Journal of Mathematics*, Vol. 52 (1930), pp. 439-500.

[5] W. Wirtinger, "Untersuchung über Thetafunctionen," Leipzig (Teubner, 1895).

[6] C. M. Huber, "On Complete Systems of Irrational Invariants of Associated Point Sets," *American Journal of Mathematics*, Vol. 49 (1927), pp. 251-267.

[7] A. B. Coble, "Associated Sets of Points," *Transactions of the American Mathematical Society*, Vol. 24 (1922), pp. 1-20.

ON THE TOPOLOGY OF ALGEBROID SINGULARITIES.

By Oscar Zariski.

The classification of algebroid singularities from a topological point of view has been the object of a thorough investigation by K. Brauner.* If the singularity consists of one branch, given by a Puiseux expansion $y = y(x)$, its intersection with the boundary of a small neighborhood, say of the 4-cell $|x| <$ const., $|y| <$ const., is projected stereographically into a knot of the ordinary space. Brauner gives a full description of these knots and he also derives the generating relations of their fundamental group.†

If the singularity consists of several branches, the above intersection consists of several linked knots.

To complete the classification of the algebroid singularities from a topological point of view it would be still necessary to prove that the different knots thus obtained are actually distinct, i. e., are not isotopic. The case of singular branches of genus 1, which give rise to knots lying on a torus, is settled by a paper by O. Schreier.‡ The purpose of the present paper is to prove the general theorem: §

* K. Brauner, "Zur Geometrie der Funktionen zweier Veränderlichen": III. Klassifikation der Singularitäten algebroider Kurven; IV. Die Verzweigungsgruppen; *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität*, Vol. 6 (1928), pp. 8-54.

† It is apparent, however, that this part of Brauner's paper, though extremely interesting in its conclusions, is unnecessarily long and complicated. E. Kähler obtains Brauner's results in a simpler manner in his paper "Über die Verzweigung einer algebraischen Funktion zweir Veränderlichen in der Umgebung einer singulären Stelle," *Mathematische Zeitschrift*, Vol. 30 (1929), pp. 188-204.

‡ O. Schreier, "Über die Gruppen $A^aB^b = 1$," *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität*, Vol. 3 (1924), pp. 167-169.

§ After having communicated this theorem at a meeting of the American Mathematical Society in March, 1932, I learnt from Professor Lefschetz that the same theorem has been proved by Werner Burau, and that Burau's paper entitled "Kennzeichnung der Schlauchknoten" is being published in the forthcoming issue of the *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität.* The polynomial $F(x)$ in section (5) of this paper is also obtained by Burau as the Alexander polynomial of the knot (See J. W. Alexander, "Topological Invariants of Knots and Links," *Transactions of the American Mathematical Society*, Vol. 30 (1928), pp. 273-306). This was to be expected since the connection between the homology characters of the k-sheeted manifolds with the knot as branch curve and the Alexander polynomial of the knot was clearly pointed out in general by Alexander in his quoted paper.

Conceding priority to Mr. Burau, I publish only an outline of my proof. I do

If two algebroid singularities, each composed of a single branch, are distinct from an algebro-geometric point of view, i. e., are either of distinct genera, or being of the same genus do not have the same pairs of characteristic numbers; they are also topologically distinct, and their fundamental groups are not isomorphic.*

The method of proof is based on the consideration of k-sheeted *cyclic Riemann manifolds* having the given singularity as locus of branch points. The idea of using these manifolds as a source of invariants has been applied by Reidemeister † in his treatment of the theory of knots and also by the author in two papers dealing with the fundamental group of plane algebraic curves.‡

It seemed advisable to the author to include in this paper another method of deriving the generating relations of the fundamental group of a singularity composed of a single branch. This method differs from the one used by Brauner in that it is algebraic and operates directly on the given singularity instead of on the corresponding knot. It can be readily extended to singularities composed of several branches.

The Fundamental Group.

1. *Preliminary remarks.* Any closed path in the residual space of the given singularity $y = y(x)$, can be deformed, one of its points remaining fixed, into a circuit lying in a "line", $x =$ const.§ If the order of the branch

---

this especially because of the additional result arrived at in the course of the proof to the effect that the Betti number R_1 of the k-sheeted Riemann manifold is the number of roots of the polynomial $F(x)$ which are also roots of the equation $x^k = 1$. Whether this is a general property of the Alexander polynomials or holds only for the particular knots connected with algebraic singularities—is a question which would be worthwhile considering.

* The characteristic numbers of a singular branch determine completely the character of the singularity in the sense of Nöther, i. e., as an aggregate of successive, infinitely near multiple points. For this and related notions of the theory of singularities see the original exposition in F. Enriques and O. Chisini, *Teoria geometrica delle equazioni e delle funzioni algebriche*, Vol. 2, book 4, Ch. I.

† "Knoten und Gruppen," *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität*, Vol. 6 (1926), pp. 7-23.

‡ "On the Linear Connection Index of the Algebraic Surfaces $z^n = f(x, y)$," *Proceedings of the National Academy of Sciences*, Vol. 15, No. 6 (1929), pp. 494-501; "On the Irregularity of Cyclic Multiple Planes," *Annals of Mathematics*, series 2, Vol. 32, No. 3 (1931), pp. 485-511.

§ See S. Lefschetz, *L'Analysis Situs et la géometrie algébrique*, p. 33. See also my paper "On the Problem of Existence of Algebraic functions of Two Variables Possessing a Given Branch Curve," *American Journal of Mathematics*, Vol. 51, No. 2 (April, 1929), where the method outlined below has been applied to the similar problem at large, involving algebraic curves instead of the neighborhood of an algebroid singularity.

is n, that circuit is a product of loops $g_0, g_1, \cdots, g_{n-1}$, each surrounding one of the points $y_0, y_1, \cdots, y_{n-1}$, in which that line intersects the branch. Hence the loops g_i can be taken as generators of the fundamental group of the given singularity. As x makes one turn around the origin $x = 0$, (the fixed initial point of the considered circuits being the point at infinity on the y-axis), the original set of non-intersecting loops g_i is continuously deformed into a new set of non-intersecting loops g'_i. We obtain all the generating relations of the fundamental group by expressing the loops g'_i in terms of the loops g_i, and by putting $g'_i = g_i (i = 0, 1, 2, \cdots, n-1)$.*

Let
$$y = \sum_{\nu=1}^{\infty} a_\nu x^{\alpha_\nu/n},$$

be the Puiseux expansion of the given branch, which we shall denote by Γ. We suppose that the axis $y = 0$ is tangent to the branch. Then $\alpha_1 > n$. Let $\alpha_1/n = m_1/n_1$, where m_1 and n_1 are relatively prime integers, in symbols $(m_1, n_1) = 1$. We write the exponent α_2/n of the second term of the series in the following form:

$$\alpha_2/n = m_1/n_1 + m_2/n_1 n_2,$$

where $(m_2, n_2) = 1$; and more in general, let

$$\alpha_\nu/n = m_1/n_1 + m_2/n_1 n_2 + \cdots + m_\nu/n_1 n_2 \cdots n_\nu,$$

where $(m_\nu, n_\nu) = 1$. Obviously there is only a finite number of values of ν, for which $n_\nu > 1$. If n_q is the last n_ν, which is greater than 1, then $n_1 n_2 \cdots n_q = n$. The terms $a_\nu x^{\alpha_\nu/n}$, for which $n_\nu > 1$, are called the *characteristic terms* of the Puiseux expansion, and the number p of such terms is called the *genus* of the branch.

It is well known that the type of the given singularity, from the standpoint of analytic transformations on the variables x and y, is completely determined by the characteristic terms defined above. It follows *a fortiori* (and this has also been proved directly by Brauner) that also topologically the given algebroid branch is completely determined by the exponents of the characteristic terms. We therefore replace the given series by the sum of its characteristic terms:

* The assertion that we obtain in this manner *all* the generating relations of the group follows from the fact that the relations $g'_i = g_i$ offer the sufficient conditions for the existence of functions of two variables possessing the given singularity as locus of branch points. For the proof of this last statement see F. Enriques, "Sulla costruzione delle funzioni algebriche di due variabili possedenti una data curva di diramazione," *Annali di Matematica pura ed applicata*, Ser. 4, Vol. 1 (November, 1923), pp. 185-198.

$$y = a_1 x^{m_1/n_1} + a_2 x^{m_1/n_1 + m_2/n_1 n_2} + \cdots + a_p x^{m_1/n_1 + m_2/n_1 n_2 + \cdots + m_p/n_1 n_2 \cdots n_p},$$

where $n_i > 1$, $(i = 1, 2, \cdots, p)$, and $(m_i, n_i) = 1$. Since the actual values of the coefficients are irrelevant, we shall put for conveniency $a_i = 1/2^{i-1}$.

We shall derive the fundamental group of the given branch Γ by a procedure of induction, considering first the case of branches of genus 1 and 2.

2. *The fundamental group of the branch* $y = x^{m_1/n_1}$. We let x vary on the circle $|x| = 1$ starting from the initial value $x = 1$. We mark in the plane of the complex variable y the initial values of y:

$$y_0 = 1, \; y_1 = e^{2\pi i/n_1}, \cdots, y_{n_1-1} = e^{2\pi i(n_1-1)/n_1},$$

and we fix the loops g_i so that if $m_1 = 1$, then as x makes one turn about the origin, the loops g_i should be transformed as follows:

$$g'_0 = g_1, \; g'_1 = g_2, \cdots, g'_{n_1-2} = g_{n_1-1}, \; g'_{n_1-1} = P_1^{-1} g_0 P_1, \tag{1}$$

where $P_1 = g_0 g_1 \cdots g_{n_1-1}$ is the circuit surrounding all the points y_i. We shall denote this transformation by T_1. It is obvious that P_1 is an invariant circuit: $T_1(P_1) = P_1$.

For an arbitrary m_1 the transformation of the loops obtained by letting x turn around the origin coincides with the m_1-th power of T_1. If we put

$$m_1 = h_1 n_1 + l_1,$$

where $n_1 > l_1 \geqq 0$, then $T_1^{m_1}$ can be written as follows:

$$g'_0 = P_1^{-h_1} g_{l_1} P_1^{h_1}, \; g'_1 = P_1^{-h_1} g_{l_1+1} P_1^{h_1}, \cdots, g'_{n_1-l_1-1} = P_1^{-h_1} g_{n_1-1} P_1^{h_1};$$
$$g'_{n_1-l_1} = P_1^{-h_1-1} g_0 P_1^{h_1+1}, \cdots, g'_{n_1-1} = P_1^{-h_1-1} g_{l_1-1} P_1^{h_1+1}.$$

We therefore obtain for the fundamental group, which we shall denote by G_1, the following generating relations:

$$\begin{aligned} g_{i+l_1} &= P_1^{h_1} g_i P_1^{-h_1}, & (i = 0, 1, \cdots, n_1 - l_1 - 1), \\ g_i &= P_1^{h_1+1} g_{n_1-l_1+i} P_1^{-h_1-1}, & (i = 0, 1, \cdots, l_1 - 1). \end{aligned} \tag{2}$$

Hence G_1 is generated by the two elements g_0 and P_1, since $n_1 - 1$ of the above relations can be used for expressing the remaining g_i's in terms of g_0 and P_1. By elimination we derive the following relation between g_0 and P_1:

$$g_0 = P_1^{m_1} g_0 P_1^{-m_1}, \tag{3}$$

which could have also been obtained by operating directly on g_0 by the transformation $T_1^{m_1 n_1}$. We obtain a second generating relation by replacing in $P_1 = g_0 g_1 \cdots g_{n_1-1}$ the factors $g_1, g_2, \cdots, g_{n_1-1}$ by their expressions in terms

of g_0 and P_1, derived from (2). We denote by x_1 and y_1 two positive integers such that

$$(3') \qquad x_1 m_1 = y_1 n_1 + 1,$$

and we consider the relations between the g_i's to which the transformation $T_1^{m_1 x_1}$ gives rise. These relations, which are obviously a consequence of (2), are the following:

$$(4) \qquad \begin{aligned} g_{i+1} &= P_1^{\nu_1} g_i P_1^{-\nu_1}, \qquad (i = 0, 1, \cdots, n-2), \\ g_0 &= P_1^{\nu_1+1} g_{n_1-1} P_1^{-\nu_1-1}. \end{aligned}$$

Conversely, the relations (2) follow from (3) and (4). In fact, from (4) we derive

$$\begin{aligned} g_{i+l_1} &= P_1^{\nu_1 l_1} g_i P_1^{-\nu_1 l_1}, \qquad (i = 0, 1, \cdots, n_1 - l_1 - 1), \\ g_i &= P_1^{\nu_1 l_1 + 1} g_{n_1 - l_1 + i} P_1^{-\nu_1 l_1 - 1}, \qquad (i = 0, 1, \cdots, l_1 - 1). \end{aligned}$$

Now these relations lead back to (2), since by (3) and (4) $P_1^{m_1}$ is commutative with each element g_i, and since moreover

$$(x_1 l_1 - 1) m_1 = (y_1 l_1 - h_1) n_1, \text{ i. e., } y_1 l_1 \equiv h_1 \pmod{m_1}.$$

From (4) we deduce $g_{i+1} = P_1^{i\nu_1 l_1} g_0 P_1^{-i\nu_1 l_1}$ $(i = 0, 1, \cdots, n_1 - 2)$ and substituting into $P_1 = g_0 g_1 \cdots g_{n_1-1}$, we find the required second relation between g_0 and P_1:

$$(5) \qquad (P_1^{\nu_1} g_0)^{n_1} = P_1^{m_1 x_1}.$$

Let

$$Q_1 = (P_1^{\nu_1} g_0)^{m_1} P_1^{-m_1 \nu_1}.$$

Then, by (3) and (5),

$$Q_1^{x_1} = (P_1^{\nu_1} g_0)^{m_1 x_1} P_1^{-m_1 x_1 \nu_1} = (P_1^{\nu_1} g_0)(P_1^{\nu_1} g_0)^{n_1 y_1} P_1^{-n_1 x_1 y_1},$$

or

$$(6) \qquad Q_1^{x_1} = P_1^{\nu_1} g_0,$$

and hence P_1 and Q_1 can be taken as generators of G_1. The relations (3) and (5) lead to the following relation between P_1 and Q_1:

$$(7) \qquad Q_1^{n_1} = P_1^{m_1},$$

and conversely from this relation follow (3) and (5), if we use (6) to define g_0 in terms of P_1 and Q_1. Consequently, *the group G_1 is generated by two elements P_1 and Q_1, satisfying the relation* (7).

3. *The fundamental group G_2 of a branch of genus 2*:

$$(8) \qquad y = x^{m_1/n_1} + 1/2\, x^{m_1/n_1 + m_2/n_1 n_2}.$$

Let, as before, $y_k = e^{2\pi ik/n_1}$, $(k = 0, 1, \cdots, n_1 - 1)$. Then the n_1n_2 values of y in (8), for $x = 1$, can be denoted by y_{ij}, where i runs from 0 to $n_1 - 1$ and j runs from 0 to $n_2 - 1$, in such a manner that for a fixed i the n_2 points y_{ij} lie on the circle of center y_i and of radius 1/2. We choose our notations so that arg. $(y_{i,j+1} - y_i)/(y_{ij} - y_i) = 2\pi i/n_2$, and that $y_{00} = 3/2$. It is immaterial, for the sequel, which of the points y_{ij}, for i fixed and > 0, is denoted by y_{i0}.

We shall accordingly denote the loop surrounding the point y_{ij} by g_{ij}. We choose these loops in the following manner: We first construct the loops g_i $(i = 0, 1, 2, \cdots, n_1 - 1)$, considered in the preceding case, where now g_i surrounds the points $y_{i0}, y_{i1}, \cdots, y_{i,n_2-1}$, and hence also y_i. We then choose the loops g_{ij} in such a manner that

$$g_{i0}g_{i1} \cdots g_{i,n_2-1} = g_i.$$

Let T_2 be the transformation of the loops g_{ij} affected by a complete turn of the point x around the origin. This transformation induces a transformation of the loops g_i, which coincides with the transformation T_1 considered in the preceding case. We therefore have, as a first set of generating relations of G_2, the generating relations (2) of G_1, or the relation (7) involving the reduced set of generators P_1 and Q_1. As for the new relations between the g_{ij}'s, it is not necessary to write them all explicitly, as most of them serve only to express the g_{ij}'s in terms of the g_i's and of the g_{0j}'s. For instance, the loops $g_{0,j}$ go by T_2 into transformed of the loops $g_{l_1,j}$, where l_1 is defined by (3′), and the corresponding generating relations are of the type:

$$g_{l_1,j} = A_j g_{0,j'} A_j^{-1},$$

where the elements A_j are in G_1, i. e., are expressible in terms of the g_i's, hence also in terms of P_1 and Q_1. If we express the fact that $\prod_{j=0}^{n_2-1} A_j g_{0,j'} A_j^{-1} = g_{l_1}$, we obtain a relation in which the loops $g_{0,j'}$ will necessarily combine into their product g_0, and the result will be obviously one of the relations (2), namely: $g_{l_1} = P_1^{k_1} g_0 P_1^{-k_1}$.

In a similar manner we express all the loops g_{ij} in terms of the loops $g_{0,j}$, g_i, by considering the successive powers $T_2^2, T_2^3, \cdots, T_2^{n_1-1}$ of T_2.

In order to obtain relations for this reduced set of generators of G_2, we must consider the transformation $T_2^{n_1}$, obtained by letting the variable x make n_1 complete turns around the origin. We put $x = z^{n_1}$, and hence

$$y = z^{m_1} + 1/2\, z^{m_1 + m_2/n_2},$$

and we let z make one turn around $z = 0$. If we put

(9) $$m_2 = h_2 n_2 + l_2,$$

we easily conclude that the transformation $T_2^{n_1}$ has the following form:

$$T_2^{n_1} \equiv \begin{cases} g'_{0j} = g_0^{-h_2} P_1^{-m_1} g_{0,j+l_2} P_1^{m_1} g_0^{h_2}, & (j = 0, 1, \cdots, n_2 - l_2 - 1); \\ g'_{0,n_2-l_2+j} = g_0^{-h_2-1} P_1^{-m_1} g_{0,j} P_1^{m_1} g_0^{h_2+1}, & (j = 0, 1, \cdots, l_2 - 1). \end{cases}$$

We have then, in addition to the relation (7), the following relations:

$$\text{(10)} \qquad \begin{cases} g_{0,j+l_2} = P_1^{m_1} g_0^{h_2} g_{0,j} g_0^{-h_2} P_1^{-m_1} & (j = 0, 1, \cdots, n_2 - l_2 - 1); \\ g_{0,j} = P_1^{m_1} g_0^{h_2+1} g_{0,n_2-l_2+j} g_0^{-h_2-1} P_1^{-m_1}, & (j = 0, 1, \cdots, l_2 - 1). \end{cases}$$

The relations (7) and (10) constitute a complete set of generating relations of the group G_2. Obviously a further reduction of the number of generators is possible, since the elements $g_{0,j}$, $j > 0$, can be eliminated, leaving g_{00}, P_1 and Q_1 as only generators of G_2. In order to accomplish this elimination we proceed in a manner similar to the one followed in the preceding case. Observing that by (3) g_0 is commutative with $P_1^{m_1}$, we derive from (10) the following relation:

$$\text{(11)} \qquad g_{00} = g_0^{m_2} P_1^{m_1 n_2} g_{00} (g_0^{m_2} P_1^{m_1 n_2})^{-1},$$

which could have also been derived by operating directly on g_{00} by $T_2^{n_1 n_2}$. Let x_2 and y_2 be positive integers, such that

$$\text{(12)} \qquad x_2 m_2 = y_2 n_2 + 1.$$

The transformation $T_2^{n_1 x_2}$ leads to relations which are obviously a consequence of the relations (10). These relations are:

$$\text{(13)} \qquad \begin{aligned} g_{0,j+1} &= B g_{0,j} B^{-1}, & (j = 0, 1, \cdots, n_2 - 2), \\ g_{0,0} &= B g_0 g_{0,n_2-1} g_0^{-1} B^{-1}, \end{aligned}$$

where $B = g_0^{y_2} P_1^{m_1 x_2}$. Conversely, the relations (10) follow from these relations and from the relation (11). In fact, from (13) we deduce

$$\text{(14)} \qquad \begin{aligned} g_{0,j+l_2} &= B^{l_2} g_{0,j} B^{-l_2}, & (j = 0, 1, \cdots, n_2 - l_2 - 1), \\ g_{0,j} &= B^{l_2} g_0 g_{0,n_2-l_2-1} g_0^{-1} B^{-l_2}, & (j = 0, 1, \cdots, l_2 - 1). \end{aligned}$$

Now, we have the following identity: $(y_2 l_2 - h_2) n_2 = (x_2 l_2 - 1) m_2$, and therefore $y_2 l_2 - h_2 = m_2 k$, $x_2 l_2 - 1 = n_2 k$, where k is an integer. Consequently, using throughout the fact that g_0 and $P_1^{m_1}$ are commutative, we find:

$$B^{l_2} = (g_0^{m_2} P_1^{m_1 n_2})^k g_0^{h_2} P_1^{m_1}.$$

Since, by (11), $g_0^{m_2} P_1^{m_1 n_2}$ is commutative with g_{00} and therefore also, by (13), with any of the elements $g_{0,j}$, we see that the relations (10) are a consequence of the relations (14) and (11).

Using (13) in order to express all the elements $g_{0,j}$ in terms of P_1, g_0 and g_{00}, and substituting into the relation $g_0 = g_{00} g_{01} \cdots g_{0,n_2-1}$, we find:

$$(11') \qquad (g_{00}g_0^{\nu_2}P_1^{m_1x_2})^{n_2} = g_0^{m_2x_2}P_1^{m_1n_2x_2}.$$

The relations (11) and (11′) are, together with (7), the generating relations of group G_2. It is possible to replace the relations (11) and (11′) by one relation. In fact, let us consider the following element of G_2:

$$Q_2 = (g_{00}g_0^{\nu_2}P_1^{m_1x_2})^{m_2}(g_0^{m_2}P_1^{m_1n_2})^{-\nu_2}.$$

It is then easily seen that

$$(15) \qquad Q_2^{x_2} = g_{00}g_0^{\nu_2}P_1^{m_1x_2},$$

and consequently we can use Q_2 as one of the generators of G_2, instead of g_{00}. The relations (11) and (11′) lead to the following:

$$Q_2^{n_2} = g_0^{m_2}P_1^{m_1n_2} = g_0^{m_2}Q_1^{n_1n_2},$$

and conversely (11) and (11′) follow from this relation, if we use (15) in order to define g_{00} in terms of P_1, g_0 and Q_1. We change slightly our notations, using the letters P_2 and P_3 in order to denote the elements g_0 and g_{00} respectively. With these notations we may state the final result as follows: *The group G_2 is generated by three elements P_1, Q_1 and Q_2, satisfying the following two relations:*

$$Q_1^{n_1} = P_1^{m_1}, \quad Q_2^{n_2} = P_2^{m_2}Q_1^{n_1n_2},$$

where $P_2 = Q_1^{x_1}P_1^{-\nu_1}$. The elements P_1, P_2 and P_3 have the following significance: P_3 is a loop surrounding one point, y_{00}, P_2 surrounds n_2 points y_{0j}, and P_1 surrounds all the n_1n_2 points y_{ij}.

4. *The fundamental group G_p of a branch Γ_p of genus p:*

$$(16) \qquad y = x^{m_1/n_1} + 1/2\, x^{m_1/n_1+m_2/n_1n_2} + \cdots + 1/2^{p-1}x^{m_1/n_1+m_2/n_1n_2+\cdots+m_p/n_1n_2\cdots n_p}.$$

We prove that: *The group G_p is generated by $p+1$ elements: P_1, Q_1, $Q_2, \cdots, Q_p$. The generating relations of G_p are the following:*

$$(17) \qquad Q_i^{n_i} = P_i^{m_i}Q_{i-1}^{n_{i-1}n_i}, \qquad (i = 1, 2, \cdots, p;\ Q_0 = 1),$$

where the elements P_2, $P_3, \cdots, P_p$ are defined by the following relations: *

$$(17') \qquad P_{i+1}P_i^{\nu_i}Q_{i-1}^{n_{i-1}x_i} = Q_i^{x_i}, \qquad (i = 1, 2, \cdots, p-1).$$

Here x_i and y_i are positive integers, such that

$$x_im_i = y_in_i + 1.$$

Moreover, if an element P_{p+1} is defined as follows:

$$(17'') \qquad P_{p+1}P_p^{\nu_p}Q_{p-1}^{n_{p-1}x_p} = Q_p^{x_p},$$

* See K. Brauner, *loc. cit.*

the elements $P_1, P_2, \cdots, P_{p+1}$ *have the following significance:* P_i *is a loop surrounding* $n_i n_{i+1} \cdots n_p$ *of the values of the function* y *for* $x = 1$ *(one value, if* $i = p+1$*).*

Since the theorem has already been proved for $p = 1, 2$, we prove it by induction, assuming that it is true for $p-1$, and namely for the branch Γ_{p-1}:

$$y = x^{m_1/n_1} + 1/2\, x^{m_1/n_1 + m_2/n_1 n_2} + \cdots + 1/2^{p-2} x^{m_1/n_1 + m_2/n_1 n_2 + \cdots + m_{p-1}/n_1 n_2 \cdots n_{p-1}},$$

obtained by dropping the last term in (16). We have then for the group G_{p-1} of this branch the generators $P_1, Q_1, Q_2, \cdots, Q_{p-1}$, and the generating relations:

$$Q_i^{n_i} = P_i^{m_i} Q_{i-1}^{n_{i-1} n_i}, \qquad (i = 1, \cdots, p-1), \tag{18}$$

where

$$P_{i+1} P_i^{y_i} Q_{i-1}^{n_{i-1} x_i} = Q_i^{x_i}, \qquad (i = 1, 2, \cdots, p-1). \tag{18'}$$

The elements P_i are loops in the plane, $x = 1$, of the variable y and surround $n_i n_{i+1} \cdots n_{p-1}$ points y_i of the branch Γ_{p-1}. On each circle of center y_i and radius $1/2^{p-1}$ there are n_p points y_{ij} of the branch Γ_p. The elements P_i can therefore be considered as elements of G_p, each P_i surrounding $n_i n_{i+1} \cdots n_p$ points y_{ij} of Γ_p. Let the notations be so chosen that $y_0 = 1 + 1/2 + \cdots + 1/2^{p-2}$ and $y_{00} = y_0 + 1/2^{p-1}$, and let $g_0, g_1, \cdots, g_{n_p - 1}$ be a set of loops surrounding the points $y_{00}, y_{01}, \cdots, y_{0, n_p - 1}$ respectively, and such that

$$g_0\, g_1 \cdots g_{n_p - 1} = P_p.$$

Then it follows as in the preceding case $p = 2$, that the elements g_j together with the generators $P_1, Q_1, Q_2, \cdots, Q_{p-1}$ of G_{p-1} constitute a set of generators of G_p; that the relations (18) and (18′) still hold; and that the only new generating relations of G_p are those which are obtained in the following manner: Let x turn around the origin $n_1 n_2 \cdots n_{p-1}$ times. Then each y_i will be reproduced, and the points y_{0j} will be permuted according to the following cyclic substitution:

$$(y_{00}, y_{0, l_p}\, y_{0, 2 l_p}, \cdots),$$

where

$$m_p = h_p n_p + l_p, \qquad 0 \leqq l_p < n_p.$$

There ensues a transformation T_p of the loops g_j, which furnishes the required new generating relations of G_p. It is easily seen that these relations are the following:

$$\begin{aligned} g_{j+l_p} &= A g_j A^{-1}, & (j = 0, 1, \cdots, n_p - l_p - 1), \\ g_j &= A P_p g_{n_p - l_p + j} P_p^{-1} A^{-1}, & (j = 0, 1, \cdots, l_p - 1), \end{aligned} \tag{19}$$

where

$$A = P_1^{m_1 n_2 n_3 \cdots n_{p-1}} P_2^{m_2 n_3 \cdots n_{p-1}} P_{p-1}^{m_{p-1}} P_p^{h_p}.$$

Using (18) and (18′) we can easily show that

$$A = Q_{p-1}^{n_{p-1}} P_p^{h_p}.$$

In fact, from those relations it follows in the first place that P_i *is commutative with* $Q_{i-1}^{n_{i-1}}$. Since this is obviously true for $i = 2$, let us assume that this is true for a given i and let us prove it for $i + 1$. Since by hypothesis P_i is commutative with $Q_{i-1}^{n_{i-1}}$, it follows from $Q_i^{n_i} = P_i^{m_i} Q_{i-1}^{n_{i-1} n_i}$ that $Q_i^{n_i}$ is commutative with P_i and with $Q_{i-1}^{n_{i-1}}$. From $P_{i+1} = Q_i^{x_i} P_i^{-y_i} Q_{i-1}^{-n_{i-1} x_i}$ it follows then that $Q_i^{n_i}$ is commutative with P_{i+1}. Using this result, we find:

$$P_1^{m_1 n_2 n_3 \cdots n_{p-1}} P_2^{m_2 n_3 \cdots n_{p-1}} = (Q_1^{n_1 n_2} P_2^{m_2})^{n_3 \cdots n_{p-1}} = Q_2^{n_2 n_3 \cdots n_{p-1}};$$
$$Q_2^{n_2 n_3 \cdots n_{p-1}} \quad P_3^{m_3 n_4 \cdots n_{p-1}} = (Q_2^{n_2 n_3} P_3^{m_3})^{n_4 \cdots n_{p-1}} = Q_3^{n_3 n_4 \cdots n_{p-1}};$$

etc., and finally $A = Q_{p-1}^{n_{p-1}} P_p^{h_p}$.

As in the preceding case the generators g_j, $j > 0$, can be eliminated, and the elimination leads to two new relations between the generators of the reduced set: $P_1, Q_1, Q_2, \cdots, Q_{p-1}$ and g_0. One of these relations is obtained directly from (19) and is the following:

$$g_0 = Q_{p-1}^{n_{p-1} n_p} P_p^{m_p} g_0 (Q_{p-1}^{n_{p-1} n_p} P_p^{m_p})^{-1}. \tag{20}$$

The second relation is obtained by considering instead of the relations (19) the weaker relations:

$$g_{j+1} = B g_j B^{-1}, \ (j = 0, 1, \cdots, n_p - 2), \quad g_0 = B P_p g_{n_p - 1} P_p^{-1} B^{-1}, \tag{21}$$

where

$$B = P_1^{x_p m_1 n_2 n_3 \cdots n_{p-1}} P_2^{x_p m_2 n_3 \cdots n_{p-1}} \cdots P_{p-1}^{x_p m_{p-1}} P_p^{y_p} = Q_{p-1}^{x_p n_{p-1}} P_p^{y_p}.$$

The relations (21) correspond to the transformation $T_p^{x_p}$. It is easily shown that the relations (21) together with the relation (20) lead back to (19). Using the relations (21) in order to express the elements g_j in terms of g_0 and of the remaining generators and substituting into the relation $P_p = g_0 g_1 \cdots g_{n_p - 1}$, we find the required relation:

$$(g_0 P_p^{y_p} Q_{p-1}^{x_p n_{p-1}})^{n_p} = (P_p^{m_p} Q_{p-1}^{n_{p-1} n_p})^{x_p}. \tag{20′}$$

We now put

$$Q_p = (g_0 P_p^{y_p} Q_{p-1}^{x_p n_{p-1}})^{m_p} (P_p^{m_p} Q_{p-1}^{n_{p-1} n_p})^{-y_p}.$$

Then

$$Q_p^{x_p} = g_0 P_p^{y_p} Q_{p-1} \tag{22}$$

and hence we may take Q_p as a generator to replace g_0. We find then from (20) and (20′):

$$(22') \qquad Q_p^{n_p} = g_0^{m_p} Q_{p-1}^{n_{p-1}n_p},$$

and conversely (20) and (20′) follow from this relation, if we use (22) as the definition of g_0. The last two relations (22), (22′), together with the generating relations (18), (18′) of G_{p-1}, coincide with the relations (17), (17′), (17″) given in the theorem announced above (where P_{p+1} stays for g_0).

5. *The fundamental homologies of the manifold* $z^k = \Pi(y - y_i)$. Given the Puiseux expansion $y = y(x)$ of a branch Γ_p of genus p, we consider the manifold $z^k = \prod_{i=1}^{n} (y - y_i)$, where n is the order of the branch, and k is an arbitrary positive integer. This k-sheeted cyclic Riemann manifold, which we shall call V, has the given branch Γ_p as locus of branch points. If we limit the range of values of x and y to the boundary of a 4-cell $|x| <$ const., $|y| <$ const., we obtain a closed k-sheeted Riemann manifold M, composed of k samples of the ordinary space, the k sheets being connected along a cylinder whose boundary is the knot associated with the given branch. The topological invariants of V, or of M, are topological invariants of Γ_p. We proceed to derive the fundamental homologies for the one-dimensional cycles on M using the generating relations of the fundamental group G_p of the residual space of Γ_p.

It is known that if permutability relations between the generators of G_p are added to the generating relations, then G_p, being the group of a knot, becomes the infinite free group generated by one element. In the present case this also follows immediately from the fact that G_p is also generated by the loops g_i $(i = 0, 1, 2, \cdots, n-1)$, and these loops are conjugate elements of G_p. Any of these loops can be taken as the generator of the free group. We have seen before that the element P_{p+1} is one of these loops. Hence in the above free group all the elements are powers of P_{p+1}. If we put

$$\nu_i = n_i n_{i+1} \cdots n_p, \qquad (i = 1, 2, \cdots n_p),$$
$$\nu_{p+1} = 1,$$

and if we introduce integers $\lambda_1, \lambda_2, \cdots, \lambda_p$, defined by the following recurrent relations:

$$\lambda_1 = m_1; \quad \lambda_i = m_i + \lambda_{i-1} n_{i-1} n_i,$$

we easily find that in the free group $\{P_{p+1}\}$ the elements P_i and Q_i are represented by the following powers of P_{p+1}:

$$P_i = P^{\nu_i}_{p+1}, \qquad Q_i = P_{p+1}^{\lambda_i \nu_{i+1}}, \qquad (i = 1, 2, \cdots, p).$$

It follows that the elements

$$P_s{}^j = P^s{}_{p+1} P_j P_{p+1}^{-(s+\nu_j)},$$
$$Q_s{}^j = P^s{}_{p+1} Q_j P_{p+1}^{-(s+\lambda_j \nu_{j+1})} \qquad (s = 0, 1, \cdots, k-1,\ j = 1, 2, \cdots, p)$$

produce the identical substitution on the branches of z and hence correspond to 1-cycles on V. It can also be shown that these $2pk$ cycles form a base for the cycles of their dimension.* We next write the fundamental homologies for the 1-cycles on M. In doing this we may neglect all powers of P_{p+1} in which the exponent is a multiple of k, since obviously $P^k{}_{p+1} \sim 0$ on M.

The generating relations (17) and (17′) lead to the following homologies:

$$(23) \qquad \sum_{r=0}^{n_{j-1}} Q^j{}_{i+r\lambda_j v_{j+1}} = \sum_{r=0}^{m_{j-1}} P^j{}_{i+rv_j} + \sum_{r=0}^{n_{j-1} n_j - 1} Q^{j-1}_{i+m_j \nu_j + r\lambda_{j-1}\nu_j}.$$
$$(j = 1, 2, \cdots, p;\ i = 0, 1, \cdots, k-1).$$

$$(24) \qquad \sum_{r=0}^{x_j - 1} Q^j{}_{i+r\lambda_j v_{j+1}} = P_i{}^{j+1} + \sum_{r=0}^{y_j - 1} P^j{}_{i+v_{j+1}+rv_j} + \sum_{r=0}^{n_{j-1} x_j - 1} Q^{j-1}_{i+\nu_{j+1}+y_j\nu_j+r\lambda_{j-1}\nu_j}.$$

where $P_i{}^{p+1}$ should be replaced by 0.

The relations (23) and (24) constitute a complete system of $2pk$ homologies between the $2pk$ cycles $Q_s{}^j$, $P_s{}^j$. The rank of the matrix of this system is $2pk - R_1$, where R_1 is the Betti number of M. This matrix, which we shall call A, is made up of circulants of order k. Each circulant may be associated with a polynomial of degree $k - 1$, whose coefficients are the elements of the first row of the circulant. Let $f_{2\alpha-1,\beta}$ and $f_{2\alpha-1,\beta+p}$ be the polynomials associated with the cycles $P_i{}^\beta$ and $Q_i{}^\beta$ respectively $(i = 1, 2, \cdots, k)$ in the k homologies of (23) in which $j = \alpha\, (\alpha, \beta = 1, 2, \cdots, p)$. Let $f_{2\alpha,\beta}$ and $f_{2\alpha,\beta+p}$ be the polynomials defined in a similar manner for the system (24). Then

$$f_{2\alpha-1,\alpha} = (x^{m_\alpha \nu_\alpha} - 1)/(x^{\nu_\alpha} - 1),$$
$$f_{2\alpha-1,\alpha+p-1} = x^{m_\alpha \nu_\alpha}(x^{\lambda_{\alpha-1}\nu_\alpha n_\alpha \nu_{\alpha-1}} - 1)/(x^{\lambda_{\alpha-1}\nu_\alpha} - 1),$$
$$f_{2\alpha-1,\alpha+p} = -(x^{\lambda_\alpha n_\alpha \nu_{\alpha+1}} - 1)/(x^{\lambda_\alpha \nu_{\alpha+1}} - 1);$$
$$f_{2\alpha,\alpha} = x^{\nu_{\alpha+1}}(x^{y_\alpha \nu_\alpha} - 1)/(x^{\nu_\alpha} - 1),$$
$$f_{2\alpha,\alpha+1} = 1,$$
$$f_{2\alpha,\alpha+p-1} = x^{x_\alpha m_\alpha \nu_{\alpha+1}}(x^{\lambda_{\alpha-1} x_\alpha \nu_{\alpha-1}} - 1)/(x^{\lambda_{\alpha-1}\nu_\alpha} - 1),$$
$$f_{2\alpha,\alpha+p} = -(x^{\lambda_\alpha x_\alpha \nu_{\alpha+1}} - 1)/(x^{\lambda_\alpha \nu_{\alpha+1}} - 1),$$

and f_{ij} is identically zero in every other case.

The matrix A is completely described by the matrix $B = \| f_{ij} \|$. It can be proved that if $\omega_1, \omega_2, \cdots, \omega_k$ are the k roots of the equation $x^k = 1$, and if $B(\omega_i)$ denotes the matrix obtained from B, by replacing x by ω_i, then A can be brought into a diagonal form, the elements of the main diagonal being the

* See K. Reidemeister, *loc. cit.*

matrices $B(\omega_i)$, while all the other elements are zero-matrices. It follows that the rank of the matrix A is equal to the sum of the ranks of the k matrices $B(\omega_i)$. By induction it can be easily shown that

$$\text{Det. } B = F(x) = f_1(x) f_2(x) \cdots f_p(x),$$

where

$$f_i(x) = (x^{\lambda_i \nu_i} - 1)(x^{\nu_{i+1}} - 1) / (x^{\lambda_i \nu_{i+1}} - 1)(x^{\nu_i} - 1)$$

is a polynomial in x. The proof by induction is based upon the fact that for every $p' < p$ the matrix of the elements of the first p' rows and columns of B is again of the same type as B. Namely, if we write

$$B = B(x, p, n_1, n_2, \cdots, n_p;\ m_1, m_2, \cdots, m_p)$$

in order to indicate the dependency of the matrix B on the characteristic numbers n_i and m_i and on the variable x, we find that its first p' rows and columns form a matrix $B(x^{\nu_{p'+1}}, p', n_1, n_2, \cdots, n_{p'}; m_1, m_2, \cdots, m_{p'})$. *Hence* $B(\omega_i)$ *is of rank less than* $2p$ *if and only if* ω_i *is a root of the polynomial* $F(x)$. But it can be shown that *if* $B(\omega_i)$ *is of rank* $2p - \sigma$, *then* ω_i *is precisely a* σ*-fold root of* $F(x)$, or, since all the roots of each factor $f_j(x)$ are simple, ω_i is a root of σ of the polynomials $f_j(x)$. Also this statement is proved by induction in the following manner. If $f_j(x)$ is the first of the polynomials $f_1, f_2, \cdots$, which vanishes for $x = \omega_i$, then the last $2p - 2j$ rows and columns of the matrix B form a matrix

$$B(x, 2p - 2j, n_{j+1}, n_{j+2}, \cdots, n_p;\ \lambda_{j+1}, \lambda_{j+2}, \cdots, \lambda_p) = \bar{B},$$

so that

$$\text{Det. } \bar{B} = f_{j+1}(x) f_{j+2}(x) \cdots f_p(x).$$

It is then easily proved that if $B(\omega_i)$ is of rank $2p - \sigma$, the matrix $\bar{B}(\omega_i)$ is of rank $2p - 2j - \sigma + 1$. Assuming that the statement is true for matrices $B(x, p' \cdots)$ where $p' < p$—for $p = 1$ the truth of the statement is obvious—it follows that the statement holds for the given matrix B of order $2p$.

We thus arrive at the following theorem: *The Betti number* R_1 *of the manifold* M *is equal to the number of roots of the polynomial* $F(x)$ (*each root counted to its proper multiplicity*) *which are also roots of the equation* $x^k = 1$.

It is now an easy matter to show that if two singularities are topologically equivalent their associated polynomials $F(x)$ must coincide, and hence the polynomial $F(x)$ is a topological invariant of the singularity. From this it follows immediately that also the genus of the singularity and all its characteristic numbers m_i and n_i are topological invariants.

THE JOHNS HOPKINS UNIVERSITY.

ON A THEOREM OF EDDINGTON.

By Oscar Zariski.

1. In a recent paper,* Eddington considers a set of 4-rowed complex matrices $F_1, F_2, \cdots F_n$, satisfying the following two conditions: 1) $F_i^2 = -E$, where E is the unit matrix; 2) any two matrices of the set are *anticommutative*, i. e., $F_iF_j = -F_jF_i$, if $i \neq j$. He proves that the maximal number of matrices in such a set is *five*, and that if the matrices are conditioned to be either *real* or *pure imaginary*, then *two at most* are real and the other three are pure imaginary.

Eddington's proof is algebraic. It may not be without interest to show that it is possible to derive Eddington's theorem geometrically in a simple and elegant manner by making use of known facts concerning *abelian groups of involutory collineations* in the complex or real projective space. These groups were classified by Study † for projective spaces of any dimensions. In the present case we find that the problem is connected with the so-called *collineation group of a Kummer configuration*, which is a group of order 16 all elements of which, besides the identity, are harmonic biaxial collineations.‡

2. With each matrix F_i of the set, considered as the matrix of coefficients of a linear homogeneous transformation, we may associate a collineation T_i in the complex projective space S_3. The condition $F_i^2 = -E$ says that T_i is a non-degenerate involutory collineation. The condition that any two distinct matrices of the set be anticommutative says that any two collineations T_i and T_j are permutable. Moreover, the n collineations $T_1, T_2, \cdots, T_n$ are all distinct, because if, for instance, T_i and T_j were the same, then the

* On Sets of Anticommutative Matrices," *Journal of the London Mathematical Society*, Vol. 7, Part 1 (January, 1932).

† "Gruppen zweiseitiger Kollineationen," *Göttinger Nachrichten* (1912), pp. 452-479.

‡ In a paper by M. H. A. Newman, "Note on an Algebraic Theorem of Eddington," published in the *Journal of the London Mathematical Society*, Vol. 7, Part 2 (April, 1932), the reader will find a generalization of Eddington's theorem to matrices of any order $2^n q$, q odd. Also this more general theorem could be derived, by a method similar to the one given below, from Study's classification of the maximal groups of involutory collineations in S_r, $r = 2^n q - 1$. The number of projectively distinct maximal groups and their structure depends solely on n. It is therefore obvious from a geometric point of view, that the maximal number of anticommutative $2^n q$-rowed matrices and 2^n-rowed matrices must be the same. Dr. J. Williamson, who called my attention to Eddington's theorem, was interested in its generalization and arrived at results, which agree with Newman, except for a slight difference in the possible number of real matrices in a set.

elements of the matrix F_i would be proportional to the elements of the matrix F_j, and the two matrices would be commutative, contrary to hypothesis. It follows that the collineations T_i generate an abelian group of involutary collineations, since the product of two permutable involutory collineations is again an involutory collineation. By Study (*loc. cit.*) there are only two distinct maximal groups of involutory collineations in S_3, i. e. such that every group of involutory collineations either coincides with or is a subgroup of one of them: 1) *the tetrahedral group,* 2) *the group of a Kummer configuration.* The elements of the tetrahedral group, which is of order 8, are collineations whose loci of invariant points are opposite elements of a tetrahedron, i. e., either a face and an opposite vertex or two opposite edges. It follows that the collineations T_i cannot belong to the tetrahedral group, because otherwise, assuming this tetrahedron as reference base for coördinates in S_3, the matrices F_i would be reduced simultaneously to the diagonal form and could not be anticommutative.

Having proved that the collineations T_i must belong to the group of a Kummer configuration, we recall that this group of order 16, which we shall denote by G_{16}, is obtained in the following manner. Consider a quadric surface Q in S_3, and the two families of lines on Q. Call the lines of one family-lines u, and the lines of the other family-lines v. Let $u^1, \bar{u}^1$ and $u^2, \bar{u}^2$ be two pairs of lines u harmonically related to each other. Similarly, let $v^1, \bar{v}^1$ and $v^2, \bar{v}^2$ be two harmonically related pairs of lines v. Let U_1 and U_2 be the harmonic biaxial homologies whose axes are the lines $u^1, \bar{u}^1$ and $u^2, \bar{u}^2$ respectively, and let V_1 and V_2 be the harmonic biaxial collineations whose axes are the lines $v^1, \bar{v}^1$ and $v^2, \bar{v}^2$ respectively. The four collineations U_1, U_2, V_1, V_2 generate the group G_{16}. The product $U_1U_2 = U_3$ has as its axes the pair of lines $u^3, \bar{u}^3$, which are harmonically related to both pairs $u^1, \bar{u}^1$ and $u^2, \bar{u}^2$. Similarly the axes $v^3, \bar{v}^3$ of the product $V_3 = V_1V_2$ are harmonically related to both pairs $v^1, \bar{v}^1$ and $v^2, \bar{v}^2$. The axes of the collineation U_iV_j are the diagonals of the skew quadrilateral formed by the lines $u^i, \bar{u}^i, v^j, \bar{v}^j$. The collineations of G_{16} leave invariant the quadric Q and transform the lines of each family into the lines of the same family. The collineation U_i $(i = 1, 2, 3)$ leaves each line v invariant and induces in the family of lines u an involution whose double elements are the lines $u^i, \bar{u}^i$. We shall denote this involution by π_i. Similarly let ω_i denote the involution induced in the family of lines v by the collineation V_i.

A last remark about the group G_{16}, which shall be of use in the sequel, is the following. There are 10 quadrics which are left invariant by the G_{16}, one of which is Q, and each of these quadrics is related to the G_{16} in the

same manner as Q. If the collineations of the G_{16} are real, the above quadrics are all real, and namely 9 are real ruled quadrics, while the tenth has no real points (imaginary ellipsoid).*

3. By a proper transformation of the coördinates we first reduce the equation of Q to the following canonical form:

$$x_1x_4 - x_2x_3 = 0,$$

and we also consider the parametric representation of Q:

$$x_1 = u_1v_1, \quad x_2 = u_1v_2, \quad x_3 = u_2v_1, \quad x_4 = u_2v_2,$$

where the parameters u_1, u_2 and v_1, v_2 can obviously be considered as homogeneuos coördinates of a line u and of a line v respectively. We point out immediately that if the matrices F_i are either real or pure imaginary, the collineations T_i are all real, and therefore, by the remark at the end of the preceding section, we may assume Q to be a real ruled quadric. Therefore the transformation of coördinates, which reduces the equation of Q to the above canonical form, is real.

If T is any collineation in G_{16}, and if the equations of the projectivities π and ω induced by T in the system of lines u and in the system of lines v are

$$u'_1 = a_{11}u_1 + a_{12}u_2, \quad u'_2 = a_{21}u_1 + a_{22}u_2,$$

and

$$v'_1 = b_{11}v_1 + b_{12}v_2, \quad v'_2 = b_{21}v_1 + b_{22}v_2,$$

respectively, then the equations of the collineation T are the following:

$$\begin{aligned} x'_1 &= a_{11}b_{11}x_1 + a_{11}b_{12}x_2 + a_{12}b_{11}x_3 + a_{12}b_{12}x_4, \\ x'_2 &= a_{11}b_{21}x_1 + a_{11}b_{22}x_2 + a_{12}b_{21}x_3 + a_{12}b_{22}x_4, \\ x'_3 &= a_{21}b_{11}x_1 + a_{21}b_{12}x_2 + a_{22}b_{11}x_3 + a_{22}b_{12}x_4, \\ x'_4 &= a_{21}b_{21}x_1 + a_{21}b_{22}x_2 + a_{22}b_{21}x_3 + a_{22}b_{22}x_4. \end{aligned}$$

The matrix F of the coefficients of T is nothing else than the direct product of the two matrices A and B, in symbols:

$$F = A \times B,$$

where

$$A = \begin{pmatrix} a_{11}, & a_{12} \\ a_{21}, & a_{22} \end{pmatrix}, \qquad B = \begin{pmatrix} b_{11}, & b_{12} \\ b_{21}, & b_{22} \end{pmatrix}.$$

If the matrix F is given, then the two matrices A and B are determined to within factors α and β connected by the relation: $\alpha\beta = 1$. Moreover, if

* Study, *loc. cit.*, p. 467.

$$F_1 = A_1 \times B_1, \qquad F_2 = A_2 \times B_2,$$

then

(1) $$F_1F_2 = (A_1A_2) \times (B_1B_2).$$

In particular, $F^2 = A^2 \times B^2$. Since both π and ω are involutions, we have $A^2 = \rho E$ and $B^2 = \sigma E$. We shall choose the factor α so as to have always $A^2 = E$. But then, in view of the condition $F^2 = -E$, we must have $B^2 = -E$. A and B are then determined to within a $\pm$ sign.

We have to consider 3 matrices $A: A_1, A_2, A_3$, corresponding to the 3 involutions π_1, π_2, π_3 considered above, and 3 matrices $B: B_1, B_2, B_3$, corresponding to the 3 involutions $\omega_1, \omega_2, \omega_3$. We also introduce the matrices A_0 and B_0, corresponding to the identities π_0 and ω_0. The involutions π_1, π_2, π_3 and the identity π_0 form a group, and if G_{16} is real, then of the 3 involutions π_i, all real, two are necessarily hyperbolic and one is elliptic. We may then suppose that the double elements $u^1, \bar{u}^1$; $u^2, \bar{u}^2$; $u^3, \bar{u}^3$ of the involutions π_1, π_2, π_3 correspond to the following values of the non-homogeneous parameter $u = u_1/u_2$: $0, \infty$; $+1, -1$; $+i, -i$. In view of the condition $A_i^2 = E$, we find then:

$$A_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}.$$

Making a similar supposition for the double elements of the involutions $\omega_1, \omega_2, \omega_3$ we find in view of the condition $B_i^2 = -E$, that

$$B_0 = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}, \quad B_1 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad B_2 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad B_3 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Each of the matrices $F_1, F_2, \cdots, F_n$ is the direct product of an A_i by a B_j. Let

$$F_k = A_{i_k} \times B_{j_k}, \qquad (k = 1, 2, \cdots, n)$$

where i_k and j_k are numbers of the set 0, 1, 2, 3. We can verify directly that $A_iA_j = -A_jA_i$ and $B_iB_j = -B_jB_i$ if $i \neq j$ and neither i nor j has the value 0. In the excluded cases the matrices will be of course commutative. We now use the condition that $F_kF_l = -F_lF_k$. We deduce from this condition and from (1) that the matrices A_{i_k} and A_{i_l} are commutative or anticommutative according as the matrices B_{j_k} and B_{j_l} are anticommutative or commutative. I say that if no i_k *and* no j_k is 0, then $n \leqq 3$. In fact, if all the i_k's have the same value, then any two of the matrices A_{i_k} are commutative, and therefore any two of the matrices B_{j_k} must be anticommutative, and that

is possible only if all the j_k's are distinct and different from 0, i. e., if $n \leqq 3$. If not all the i_k's are equal, let, for instance, $i_1 \neq i_2$. Then necessarily $j_1 = j_2$. For any other value of k we must also have $j_k = j_1$. In fact, if $j_k \neq j_1$, then B_{j_k} is anticommutative with $B_{j_1} = B_{j_2}$ and therefore A_{i_k} must be commutative with both A_{i_1} and A_{i_2}, which is impossible, since $i_1 \neq i_2$ and $i_k \neq 0$. Hence all the j_k's have the same value, and we conclude, as before, that $n \leqq 3$. Hence, if $n > 3$, then for some value of k we must have either $i_k = 0$ or $j_k = 0$. Suppose that, for instance, $i_1 = 0$. Then obviously $j_k \neq 0$ for any k. If now $n > 4$, then two indices i_k must be equal. If these two equal indices are not 0, then let, for instance, $i_2 = i_3 = 1$. Since any two of the matrices A_0, A_1, A_1 are commutative, the indices j_1, j_2, j_3 must be all distinct and different from 0. Let, for instance, $j_1 = 1$, $j_2 = 2$, $j_3 = 3$. For any other value of $k > 3$, we must have $j_k \neq 1$ and also $j_k \neq 2, 3$, since B_{j_k} must be either commutative or anticommutative with both B_{j_2} and B_{j_3}. Hence $j_k = 0$, which is impossible. Consequently, if $n > 4$, the two equal indices i_k must be equal to 0. Let $i_1 = i_2 = 0$. Then j_1 and j_2 must be distinct and again no j_k can be equal to 0. For any $k > 2$, we must have $j_k \neq j_1, j_2$. Hence $j_3 = j_4 = j_5 = \cdots = j_n$, and for this reason the indices $i_3, i_4, i_5, \cdots, i_n$ must be all distinct and different from 0. We conclude that $n = 5$ and, for instance, that $i_3 = 1$, $i_4 = 2$, $i_5 = 3$, and that the matrices F_k are the following:

$$F_1 = A_0 \times B_{j_1},\ F_2 = A_0 \times B_{j_2},\ F_3 = A_1 \times B_{j_3},\ F_4 = A_2 \times B_{j_3},\ F_5 = A_3 \times B_{j_3},$$

where the numbers j_1, j_2, j_3 coincide with the numbers 1, 2, 3, in some order. This proves the first part of Eddington's theorem.

From the explicit expression of the matrices A_i and B_i given above, we see that if $j_3 = 3$ then the matrices F_3 and F_4 are real and the others are pure imaginary, whereas if $j_3 \neq 3$, say if $j_1 = 1$, $j_2 = 3$, $j_3 = 2$, then F_2 and F_5 are real and the others are pure imaginary. This proves the second part of Eddington's theorem.

THE JOHNS HOPKINS UNIVERSITY.

REGULAR LINEAR SYSTEMS OF CURVES WITH THE SINGULARITIES OF A GIVEN CURVE AS BASE POINTS.

By MARGUERITE LEHR.

Summary. An irreducible algebraic curve of order $m, f(x,y)=0$, in the plane $\pi(z=0)$ is given with singularities of a specified kind. The surface $F, z^n=f(x,y)$ where n is a positive integer, gives rise to adjoint surfaces ϕ_ν, and to complete linear systems of surfaces $\phi_\nu - d\pi$ which together with the plane π taken d times form adjoints ϕ_ν. These systems of surfaces cut the plane π in systems of curves $|C^d|$, shown to be complete. The irregularity of the surface F can be expressed in terms of the superabundances of some of the systems $|C^d|$, hence application of the theorem * that F is regular if n is a power of a prime shows the regularity of the systems of plane curves involved. For $f(x,y)=0$ having only nodes and cusps, Zariski † obtains in this manner regular systems of curves with the cusps of f as base points. In this paper we consider f possessing either of two types of singularities:— ordinary k-fold points and k-fold points each of which is the origin of one branch of order k. The resulting theorems are analogous to the well-known theorem: Curves of order $m-3$ with $(k-1)$-fold points at the k-fold points of f form a regular system; we prove that curves of order $m-3-j$ with $(k-i)$-fold points $(i=1,2,\cdots,k-1)$ at the k-fold points of f form regular systems for j having a certain range of values. Thus restrictions are imposed on the number and position of specified singularities for f. Further, since the surface F is regular for all positive integral values of n when $f(x,y)$ has a cyclic fundamental group with respect to its carrying complex plane,† the connection between the irregularity of F and the superabundances of the linear systems described enables us to state sufficient conditions that $f(x,y)=0$ have a non-cyclic fundamental group; these conditions show that there exists a connection between the position of the singular points other than nodes and the structure of the fundamental group.

The adjoints $\phi_\nu(x,y,z)=0$ are conditioned in general at all singularities of $z^n=f(x,y)$. This surface F has isolated singularities only at the singular

* O. Zariski, "On the Linear Connection Index of the Algebraic Surfaces $z^n=f(x,y)$," *Proceedings of the National Academy of Sciences*, Vol. 15 (No. 6, June, 1929).

† O. Zariski, "On the Irregularity of Cyclic Multiple Planes," *Annals of Mathematics*, 2nd ser., Vol. 32 (No. 3, June, 1931).

points of f when $n = m$; if $n > m$ the line at infinity in the plane π is multiple on F, the only hypermultiple points on it being intersections with $f(x, y) = 0$. This line may be considered in generic position with respect to f, having m distinct intersections with f. In the paper quoted,* Zariski gives a necessary and sufficient condition imposed on adjoints ϕ_ν by the multiple line, and shows that the hypermultiple points on it impose no additional conditions. We shall determine the conditions imposed on ϕ_ν at isolated singularities of F due to various types of singular points on $f(x, y) = 0$. These together with the condition mentioned will completely describe the adjoint surfaces ϕ_ν. For the sake of clarity, the case of f possessing only ordinary k-fold points will be handled fully first; the procedure can then be modified for other conditions on f.

I. f HAVING ORDINARY k-FOLD POINTS ONLY.

1. *Necessary and sufficient conditions on ϕ_ν at O.* Let O be an ordinary k-fold point of $f(x, y) = 0$, which we may suppose to be at the origin. If $\phi_\nu(x, y, z) = 0$ is to be adjoint to $z^n = f(x, y)$, the double integral

$$\int\int \frac{\phi_\nu(x, y, z)}{z^{n-1}}\, dx dy$$

must be finite for every analytical 2-cell containing $x = y = 0$. Since the required condition is of differential character, we may replace $f(x, y)$ by the product $\Pi(y - a_i x) = (y - a_1 x)(y - a_2 x) \cdots (y - a_k x)$. Introduce two new independent variables u, t by the relations

$$(1) \qquad x = u^n, \quad y = u^n t, \quad z^n = u^{kn}\Pi(t - a_i) = u^{kn}\Pi(t).$$

Let $c_{\alpha\beta\gamma}x^\alpha y^\beta z^\gamma$ be the general term of the polynomial $\phi_\nu(x, y, z)$. The transformed double integral is then the sum of integrals of the form

$$\int\int \frac{u^{\alpha n+\beta n+\gamma k+2n-1} t^\beta \{P(t)\}^{\gamma/n}}{u^{k(n-1)}\{P(t)\}^{n-1/n}}\, du dt,$$

which must be finite for $u = 0$, t arbitrary. So far as t is concerned, the integral is finite for all finite values of t including those for which $P(t) = 0$; the infinite value for t can be investigated by using $x = u^n t, y = u^n$, which leads to the same form of integral. Hence a *necessary and sufficient condition that the double integral remain finite is*

$$(2) \qquad (\alpha + \beta - k + 2)n + k(\gamma + 1) > 0.$$

This condition requires $\alpha + \beta + \gamma \geqq k - 2$ and is satisfied by

* O. Zariski, "On the Irregularity of Cyclic Multiple Planes," *Annals of Mathematics*, 2nd ser., Vol. 32 (No. 3, June, 1931).

$$\alpha+\beta=k-2,\ \gamma=0;$$

so every ϕ_ν has at least a $(k-2)$-fold point at O, and the generic ϕ_ν cuts the plane $\pi(z=0)$ in a curve C of order $\nu-\sigma_0$ (where σ_0 accounts for the line at infinity occurring in the intersection), having an ordinary $(k-2)$-fold point at O. On the other hand, consider any curve of order $\nu-\sigma_0$ with the points * $O\,(k-2)$-fold in the plane π, and project it from the point at infinity on the z-axis. The resulting cone satisfies the adjoint condition at O given by (2). Since y^{σ_0} satisfies the conditions imposed † by the multiple line at infinity in the plane π, the degenerate surface made up of y^{σ_0} and the cone is an adjoint ϕ_ν. Every such curve can be cut out thus by an adjoint surface, so the system of curves cut out on π by $|\,\phi_\nu\,|$ is complete.

2. *Nature of the curves* $|\,C^d\,|$. Consider all surfaces $\phi_{\nu-d}$ which with the plane π taken d times give ϕ_ν adjoint. They form a complete linear system which we denote by $|\,\phi_\nu-d\pi\,|$, and this system cuts the plane π in a system of curves $|\,C^d\,|$. We wish to describe these curves for arbitrary values of d. Surfaces $\phi_\nu-d\pi$ are those ϕ_ν satisfying (2), in whose terms $\gamma\geq d$. When they are cut by $z=0$, the curves C^d are given by terms in ϕ_ν in which $\gamma=d$; hence in the equation of $C(x,y)=0$, the exponents of $x^\alpha y^\beta$ satisfy the relation

$$(\alpha+\beta-k+2)n+k(d+1)>0. \tag{3}$$

The surfaces ϕ_ν cut π in a $(k-2)$-fold point at O, and we know that there are in addition, in general, infinitely near multiple points to O, so removing $z=0$ as a factor from ϕ_ν will not necessarily reduce the multiplicity of O on the curve of section. From (3) we see that if terms $\alpha+\beta=k-3$ are to be present in the equation of C, then

$$d\geq[n/k]$$

where $[n/k]$ means the integral part of n/k. So the systems

$$|\,\phi_\nu\,|,\ |\,\phi_\nu-\pi\,|,\ \cdots\ |\,\phi_\nu-\{[n/k]-1\}\pi\,|$$

cut $z=0$ *in curves with* $(k-2)$*-fold points at the* k*-fold points of* f. These points are ordinary $(k-2)$-fold points since any terms for which

$$\alpha+\beta=k-2$$

satisfy condition (3). The curves C^d cut out by $|\,\phi_\nu-d\pi\,|$ can and will

* The only singularities of f are ordinary multiple points of one order k; we shall say "the points O" meaning that set of points.

† O. Zariski, "On the irregularity, etc.," *loc. cit.*

4

have O as a point of multiplicity $k-2-i$ if and only if d satisfies the relation

$$-in+kd+k>0 \text{ or } d \geqq [in/k]. \qquad (i=0,1,\cdots,k-2).$$

The curves C^d, then, cut out on π by $|\phi_\nu - d\pi|$ have the ordinary k-fold points of f as ordinary $(k-2-i)$-fold points for values of d given by

$$(4) \qquad [in/k] \leqq d < [(i+1)n/k]. \qquad (i=0,1,\cdots,k-2).$$

Finally for $d \geqq [(k-2)n/k]$, the systems $|\phi_\nu - d\pi|$ are unconditioned at the ordinary k-fold points of $f(x,y)=0$.

The order of the curves $|C^d|$ is $\nu-d-\sigma_d$, where σ_d accounts for the multiple line at infinity. Write

$$(5) \qquad n-m=hm-m_1, \qquad h>0,\ 0 \leqq m_1 < m.$$

For values $d \leqq n-2$, $\nu-d-\sigma_d=\nu-n+m+1-\mu$, where μ is the positive integer defined * by

$$(6) \quad (h+1)\mu+[\mu m_1/m]-2 < d \leqq (h+1)(\mu+1)-[(\mu+1)m_1/m]$$

For $d>n-2$, the surfaces are not only unconditioned at the multiple line, but they are also unconditioned at the k-fold points of $f(x,y)=0$, since $[(k-2)n/k] \leqq n-2$ for all positive integral k and n. For $d \geqq n-2$, the order of $|C^d|$ is $\nu-d$. *The systems*

$$|C^0|,\ |C^1|,\cdots,\ |C^{n-2}|$$

therefore distribute themselves into sets of consecutive systems all of order $\nu-n+m+1-\mu$ where μ runs from 0 to $m-1$. There are $h+\epsilon_\mu$ systems in each set, where

$$(7) \qquad \begin{array}{l} \epsilon_\mu=1 \text{ if } [\mu m_1/m]=[(\mu+1)m_1/m] \text{ and } \mu \neq 0, \\ \epsilon_\mu=0 \text{ if } [\mu m_1/m]=[(\mu+1)m_1/m]-1, \text{ or } \mu=0. \end{array}$$

In particular, the first h systems are of order $\nu-n+m+1$, and the last h of order $\nu-n+2$.

3. *Completeness of* $|C^d|$. Consider any curve of order $\nu-d-\sigma_d$ lying in the plane π and having points of multiplicity

$$k-2-i \qquad (i=0,1,\cdots,k-2)$$

at an ordinary k-fold point of f. The cone obtained by projecting this curve

* O. Zariski, "On the irregularity of cyclic multiple planes," *Annals of Mathematics*, 2nd ser., Vol. 32 (No. 3, June, 1931).

from the point at infinity on the z-axis together with the plane π taken d times satisfies condition (2); it is a surface, therefore, which satisfies the conditions imposed on adjoints at isolated singularities of F. As previously stated, $y^{\sigma d}$ will satisfy the conditions imposed by the multiple line, hence the cone together with σ_d times y and d times π will give an adjoint ϕ_ν. The cone plus σ_d times y is then a $\phi_\nu - d\pi$, and since every C^d of the type described can be cut out by such a surface, the systems $| C^d |$ cut out by $| \phi_\nu - d\pi |$ and described by the relations (4) and (6) are complete.

4. *Virtual and effective dimensions of* $| \phi_\nu - d\pi |$. If a system $| \phi_\lambda |$, linear and complete, cuts a plane π in a complete system of curves $C_{\lambda-\sigma}$, and if we denote by $| \phi_\lambda - \pi |$ the complete system of surfaces $\phi_{\lambda-1}$ which with π make a ϕ_λ, then for π a generic plane it is well-known that

$$(r_\lambda - r'_\lambda) = (r_{\lambda-1} - r'_{\lambda-1}) + s_\lambda,$$

where r_λ, r'_λ are respectively the effective and virtual dimensions of $| \phi_\lambda |$; $r_{\lambda-1}$, $r'_{\lambda-1}$ respectively those of $| \phi_\lambda - \pi |$, and s_λ is the superabundance of the complete system of curves C_λ cut out on π (σ being zero). In the case of the systems $| \phi_\nu - d\pi |$, however, the plane is not generic; it goes through the multiple line and all isolated singularities. It is true, nevertheless, that $| \phi_\nu - d\pi |$ cuts π in a complete system $| C^d |$, and we shall show that under this condition

$$r_d - r'_d = r_{d+1} - r'_{d+1} + s_d$$

where r_d, r'_d are respectively the effective and virtual dimensions of $| \phi_\nu - d\pi |$ and s_d is the superabundance of C^d. If we denote by ρ_d, ρ'_d the effective and virtual dimensions respectively of the complete system $| C^d |$, we can write

$$r_d - r'_{d+1} = \rho_d + 1.$$

Since
$$r'_d = \binom{\nu - d + 3}{3} - 1 - K_d(\nu - d) - K'_d,$$

$$r'_{d+1} = \binom{\nu - d + 2}{3} - 1 - K_{d+1}(\nu - d) - K'_{d+1},$$

K_d, K'_d, K_{d+1}, K'_{d+1} being the postulation constants independent of ν, for any fixed d, or, in other words, algebraic functions of d alone, we obtain

$$(8) \qquad r'_d - r'_{d+1} = \binom{\nu - d + 2}{2} - (\nu - d)(K_d - K_{d+1}) - (K'_d - K'_{d+1})$$

valid for all non-negative integral values of ν and of d. Now when $\nu - d$ is sufficiently high, $r_d = r'_d$ and $r_{d+1} = r'_{d+1}$; it follows that

$$r'_d - r'_{d+1} = r_d - r_{d+1} - \binom{\nu - d - \sigma_d + 2}{2} - K_d'' + s_d$$

where K_d'' is the postulation in the plane on the curves C^d. Since σ_d is the multiplicity of the line at infinity in π on $\phi_\nu - d\pi$, σ_d is independent of ν; for high values of ν the order of the curves C^d increases and s_d becomes zero. For $\nu - d$ sufficiently high, then,

$$(9) \qquad r'_d - r'_{d+1} = \rho'_d + 1.$$

The two expressions in (8), (9) are equivalent for all non-negative integral values of ν and of d for which $\nu - d$ is sufficiently high, *i. e.*,

$$\binom{\nu - d + 2}{2} - (\nu - d)(K_d - K_{d+1}) - (K'_d - K'_{d+1}) = \binom{\nu - d - \sigma_d + 2}{2} - K_d'' + s_d,$$

the K's and σ_d being independent of ν. Fix d at a value d_0. The two polynomials in ν are identical for all ν sufficiently high, hence they are identical polynomials and their coefficients (algebraic functions of d_0) are the same. These relations between the coefficients hold for all non-negative integral values of d, hence the coefficients are identical and the two expressions are equivalent for all values of ν, d. Consequently, so long as $|\phi_\nu - d\pi|$ cuts π in a complete system

$$r'_d - r'_{d+1} = \rho'_d + 1$$

regardless of whether π is generic or not. Since

$$r_d - r_{d+1} = \rho'_d + s_d + 1,$$

we obtain the desired relation

$$(10) \qquad r_d - r'_d = r_{d+1} - r'_{d+1} + s_d.$$

5. *Irregularity of F in terms of s_d.* Adding the relations (10) for $d = 0, 1, \cdots, [(k-2)n/k] - 1$, we find

$$(11) \qquad r_0 - r'_0 = s_0 + s_1 + \cdots + s_{[(k-2)n/k]-1} + (r_{[(k-2)n/k]} - r'_{[(k-2)n/k]}).$$

This relation holds even if the systems $|\phi_\nu - d\pi|$ or $|C^d|$ cease to exist, if we adopt the convention that the effective dimension of a non-existent system is -1.

For $\nu \geq n - 4$, $r_{[(k-2)n/k]} - r'_{[(k-2)n/k]}$ can be shown to be zero. Consider the subadjoint surfaces $\bar{\phi}_\nu$; they are conditioned at the multiple line as are ϕ_ν, and are not conditioned at the isolated singularities, so

$$| \phi_\nu - [(k-2)n/k]\pi | \text{ and } | \bar{\phi}_\nu - [(k-2)n/k]\pi |$$

are coincident systems. Now all systems $| \bar{\phi}_\nu - d\pi |$ cut out complete systems $| \bar{C}^d |$ by even stronger reasoning than for $| \phi_\nu - d\pi |$, so

$$\bar{r}_d - \bar{r}'_d = \bar{r}_{d+1} - \bar{r}'_{d+1} + \bar{s}_d.$$

But $| \bar{C}^d |$ is virtually and effectively the system all curves of order $\nu - d - \sigma_d$, of the same order as $| C^d |$, but unconditioned at isolated singularities, so $\bar{s}_d = 0$, if the order is $\geqq -2$. It follows that

$$\bar{r}_0 - \bar{r}'_0 = \bar{r}_{d+1} - \bar{r}'_{d+1},$$

for any d such that $\nu - d - \sigma_d \geqq -2$. When $d \leqq [(k-2)n/k] - 1$, the order of C^d is $\geqq \nu - n + 2$, so if $\nu \geqq n - 4$, curves C^d are all of order greater than or equal to -2. When

$$d = [(k-2)n/k] - 1,$$

then

$$\bar{r}_0 - \bar{r}'_0 = r_{[(k-2)n/k]} - r'_{[(k-2)n/k]}.$$

Since $| \phi_\nu - [(k-2)n/k]\pi |$ is identical with $| \bar{\phi}_\nu - [(k-2)n/k]\pi |$. Now by a theorem of Castelnuovo,* $\bar{r}_0 - \bar{r}'_0$ will be zero if $| \phi_{\nu+1} |, | \phi_{\nu+2} |, \cdots$ cut out on a generic plane complete and regular linear systems of curves. If $\nu \geqq n - 4$, these systems cut out curves of order $\geqq n - 3$, adjoint to the plane section of F, and it is well-known that these curves form regular systems. They are clearly complete, for any curve $C_{\nu'}$ in a generic plane ω, adjoint to the section of F made by ω, can be projected from the point at infinity on the x-axis giving a cone which is a subadjoint surface $\bar{\phi}_{\nu'}$ of F. This establishes the fact that $r_{[(k-2)n/k]} - r'_{[(k-2)n/k]} = 0$ if $\nu \geqq n - 4$. Restating (11), then, for $\nu = n - 4$, we obtain the following expression for the irregularity of F:

$$q = s_0 + s_1 + \cdots + s_{[(k-2)n/k]-1}. \tag{12}$$

The systems $| C^d |$ involved all occur in the set $| C^0 |, | C^1 |, \cdots, | C^{n-2} |$ described in § 2, and s_d is the superabundance of the system $| C^d |$ for which the order is given by (6) and (7), and the multiplicity at the ordinary k-fold points of f is given by (4). Hence *the irregularity q of F is equal to the sum of the superabundances of the first $[(k-2)n/k]$ systems $| C^d |$ described.*

6. *Regular systems* $| C^d |$. If f is irreducible, curves of order $m - 3$

* G. Castelnuovo, "Alcune proprietà fondamentali dei sistemi lineari di curve tracciate ecc . . . ," *Annali di Matematica pura ed applicata*, ser. 2, Vol. 25 (1897).

with k-fold points of f as $(k-1)$-fold points form a regular system, so s_0 and possibly certain others succeeding it are zero. Suppose that $m-3-j$ is the maximum order for which the base points O do not give rise to a regular system, *i. e.*, curves of order $m-3-\mu$ with points $O(k-2)$-fold are regular and s_μ^{k-2} is zero if $\mu < j$, but $s_\mu^{k-2} > 0$ for $\mu = j$. The irregularity of F involves such superabundances in the first $[n/k]$ terms, hence if the order $m-3-j$ is such that the number of systems preceding it in the sequence is less than $[n/k]$, a non-vanishing superabundance will appear in (12) and the surface F will be irregular. The first system of order $m-3-j$ is preceded by

$$\sum_{i=0}^{j-1} h + \epsilon_i = j(h+1) - 1 - [jm_1/m]$$

terms. Denote the difference between this and $[n/k]$ by λ.

$$\lambda = j(h+1) - 1 - [jm_1/m] - [n/k]. \tag{13}$$

Then if $\lambda < 0$, $q > 0$ and F is irregular. Write

$$n = \rho k + \sigma, \qquad \rho > 0,\ 0 \leq \sigma < k.$$

Recalling that $n = (h+1)m - m_1$, we find from (5) and (13)

$$k\lambda = (h+1)(kj-m) - k - k[jm_1/m] + m_1 + \sigma. \tag{14}$$

As n becomes indefinitely large, h increases indefinitely, but all other elements in λ are either fixed or have an upper limit set by k or by m, so λ takes its sign from $kj - m$. If $kj < m$, λ would be negative for all values of n sufficiently high. We know, however, that if n is a power of a prime, F is regular and λ cannot be negative for such values. Hence kj must be greater than or equal to m. We state the geometric result: If $f(x, y) = 0$ be irreducible of order m, with ordinary k-fold points O, and if μ be the maximum integer such that $k\mu < m$, then the complete systems of curves of orders $m-3$, $m-4, \cdots, m-3-\mu$ with the points $O(k-2)$-fold are regular.

Next, let $m-3-j$ be the maximum order for which the points $O(k-3)$-fold form the base of a superabundant system. For each system $|C^d|$, $d = 0, 1, \cdots, [2n/k]-1$, set up a system of the same order with $O(k-3)$-fold; the superabundance s_d^{k-3} of such a system is not greater than the superabundance s_d occurring in q. If in the sequence $|C^0|, \cdots, |C^{n-2}|$, then, the first system of order $m-3-j$ is preceded by less than $[2n/k]$ others, a non-vanishing superabundance will occur in (12) and F will be irregular. Writing

$$\lambda = j(h+1) - 1 - [jm_1/m] - [2n/k],$$

we find, precisely as above, $kj \geq 2m$. Hence for $k\mu < 2m$, the systems of order $m - 3 - \mu$ with points $O(k - 3)$-fold are regular. The general theorem may be derived by application of this method, using (4) to give the number of systems $| C^d |$ involved when d is set.

THEOREM. *If $f(x, y) = 0$ be irreducible of order m, with only ordinary k-fold points O, and if μ be the maximum integer such that $k\mu < (i + 1)m$, $(i = 0, 1, \cdots, k - 2)$ then all complete systems of curves of orders $m - 3$, $m - 4, \cdots, m - 3 - \mu$ with the points $O(k - 2 - i)$-fold are regular.*

7. *Conditions on the k-fold points of $f(x, y)$.* The theorem just proved excludes the possibility of an irreducible curve having so many k-fold points that the virtual dimension of any of the systems described becomes less than -1. It also precludes the ordinary k-fold points of an irreducible curve from having such special positions in the plane that they do not impose independent conditions on the curves of the orders stated, when used to the multiplicity described in the base set. An example will show how this condition operates. A curve of order 14 can have 26 triple points; by the theorem just proved, curves of order 7 passing simply through these points must form a regular system. (Note that for f irreducible, a repeated application of the theorem on adjoints of order $m - 3$ will give curves of order eight through the triple points forming a regular system; the present theorem lowers the degree here.) Fix sixteen of the eighteen intersections of a cubic and a sextic. Any C_7 through the fixed points cuts the sextic in addition in 26 points which cannot be triple points of an irreducible C_{14}, as the following argument shows. Consider all C_7's through the 26 points; the characteristic series cut out on a generic C_7 of the system is a complete g_{23}, and the system is regular or superabundant according as the series is non-special or special.* Among the groups of the g_{23} is one made up of the sixteen points fixed on the sextic and an arbitrary line section of C_7. This group is also cut out on C_7 by a quartic made up of the given cubic (with its additional five points of intersection with C_7 fixed) and an arbitrary line, hence this group is contained in the complete series cut out by canonical adjoints through five fixed points. Since the genus of a non-singular C_7 is 15, the canonical series is g^{14}_{28}; the complete series determined by the group of 23 points is therefore g^{9}_{23}, special; hence the system of curves C_7 is superabundant. This contradicts the theorem; the group of 26 points described cannot be triple points of an irreducible C_{14}.

---

* Guido Castelnuovo, "Ricerche generali sopra i sistemi lineari di curve piane," *Torino Memorie*, Ser. 2, Vol. 42 (1891).

II. f Having at each k-fold Point one Branch of Order k.

8. *Necessary and sufficient condition on ϕ_ν at O.* Again suppose O, k-fold, to be at the origin, and consider first O as the origin of a single branch of order k and class 1. The condition being of differential character, $f(x, y)$ may be replaced by $y^k + x^{k+1}$. If $\phi_\nu(x, y, z)$ is to be adjoint to $z^n = f(x, y)$, the double integral

$$\int\int \frac{\phi_\nu(x, y, z)}{z^{n-1}}\, dxdy$$

must be finite for every analytical two-cell containing $x = y = 0$. Introduce two new variables u, t by the relations:

$$x = u^{nk},\ y = u^{n(k+1)}t.$$

By considerations precisely like those of part I, we find that a *necessary and sufficient condition that the double integral remain finite is*

(15) $\qquad kn(\alpha + \beta - k + 1) + n(\beta + 1) + k(k + 1)(\gamma + 1) > 0.$

Clearly, if $\alpha + \beta \geqq k - 1$, the condition is satisfied; further, if $n \geqq k(k + 1)$ no terms for which $\alpha + \beta + \gamma < k - 1$ can appear. So the generic ϕ_ν cuts the plane π in a curve with O as $(k - 1)$-fold point. As in part I, we see that any curve in π with O as ordinary $(k - 1)$-fold point can be projected from the point at infinity on the z-axis, giving a cone which satisfies the conditions imposed on adjoints by the point O, and the system cut out on π by $| \phi_\nu |$ is complete.

9. *Nature of the curves $| C^d |$.* We consider the systems of curves $| C^d |$ cut out on π by the systems of surfaces $| \phi_\nu - d\pi |$. The curves $| C^d |$ are given in the plane $z = 0$ by terms $x^\alpha y^\beta z^d$ of ϕ_ν, hence terms in the equation of $C(x, y) = \sum c_{\alpha\beta} x^\alpha y^\beta$ satisfy the relation

(16) $\qquad kn(\alpha + \beta - k + 1) + n(\beta + 1) + (d + 1)k(k + 1) > 0.$

The system $| C^0 |$ has points $O\,(k - 1)$-fold with no further restriction. If terms $\alpha + \beta = k - 2$ are to satisfy the condition, the minimum value of d is given by the maximum $\beta = k - 2$, or

$$d \geqq [n/k(k + 1)].$$

Hence the systems $| \phi_\nu |$, $| \phi_\nu - \pi |, \cdots, | \phi_\nu - \{[n/k(k + 1)] - 1\}\pi |$ cut the plane in curves with points O as $(k - 1)$-fold points. Stating the general case, then, the curves cut out by $| \phi_\nu - d\pi |$ can and will have O as multiple point of order $k - 1 - i$ if and only if d satisfies the relation

$$(17) \qquad -ikn + n(\beta + 1) + k(k+1)(d+1) > 0, \quad (i = 0, 1, \cdots, k-1).$$

Again the minimum d satisfying this condition is given by the maximum $\beta = k - 1 - i$, so

$$d \geqq [\{(i-1)k + i\}n/k(k+1)].$$

The curves C^d cut out by $|\phi_\nu - d\pi|$ have O as $(k-1-i)$-fold point for values of d given by

$$(18) \qquad \left[\frac{(i-1)k+i}{k(k+1)}n\right] \leqq d < \left[\frac{(ik+i+1)}{k(k+1)}n\right].$$

In particular, for $d \geqq [(k^2-k-1)n/k(k+1)]$, the curves are unconditioned at points O.

In addition, the singularity at O on $|C^d|$ will in general be specialized, having a certain number of tangents coinciding with the tangent to f at O. Actually the condition (17) gives

$$(19) \qquad d \geqq \left[\frac{(ik-1-\beta)n}{k(k+1)}\right]$$

for $O(k-1-i)$-fold or less. The maximum value of β gives the minimum value of d stated in (18); for this, $|C^d|$ has all tangents at O coincident with the tangent to f at O. As β decreases, d definitely increases until we reach $d = [(ik-1)n/k(k+1)]$ which gives $O(k-1-i)$-fold with all tangents freed from coincidence with the tangent to f at O. From this value to $d = [(ik+i+1)n/k(k+1)]$ the curves $|C^d|$ have the same behavior at O. The order of the curves is given in part I. For n sufficiently large, i. e., if

$$[(k^2-k-1)n/k(k+1)] \leqq n-2$$

the surfaces $|\phi_\nu - d\pi|$ for $d \geqq n-2$ are unconditioned at the isolated points as before. Hence *the systems $|C^0|, \cdots, |C^{n-2}|$ distribute themselves into sets of consecutive systems all of orders $\nu - n + m + 1 - \mu$, where μ runs from 0 to $m-1$. Any given $|C^d|$ has its multiplicity $k-1-i$ at points O determined by* (18), *and it has at each such point τ tangents coinciding with the tangent to f there, where τ is the minimum integer satisfying the relation*

$$-ikn + n(\tau+1) + (d+1)k(k+1) > 0.$$

10. *Irregularity of F in terms of s_d.* The systems $|C^d|$ are complete, by considerations analogous to those in part I, hence the results of paragraph 4 there apply; further, for n sufficiently high, the surfaces $|\phi_\nu - d\pi|$ for $d = [(k^2-k-1)n/k(k+1)] \equiv \delta$ are conditioned at the multiple line

only, so the argument of paragraph 5 on subadjoint surfaces can be applied. Summing (10) then for values $d = 0, 1, \cdots, [(k^2 - k - 1)n/k(k+1)]$, we obtain

$$(20) \qquad r_0 - r'_0 = s_0 + s_1 + \cdots + s_{\delta-1} + r_\delta - r'_\delta.$$

For $\nu \geqq n - 4$, therefore,

$$(21) \qquad q = s_0 + s_1 + \cdots + s_{\delta-1}$$

where the s_d are superabundances of the systems $| C^d |$ described in 9.

11. *Regular Systems.* If f is irreducible, curves of order $m - 3$ with the points O as $(k-1)$-fold points form a regular system, so s_0 and possibly others are zero. Suppose that $m - 3 - j$ is the maximum order for which the base points $O(k-1)$-fold give rise to a superabundant system. The first system of order $m - 3 - j$ in the set $| C^0 | \cdots | C^{n-2} |$ is preceded by $j(h+1) - 1 - [jm_1/m]$ systems. If this is less than $[n/k(k+1)]$, a non-vanishing superabundance will occur in (21) and F will be irregular. If $k(k+1)j < m$, then, F will be irregular for all n sufficiently high, as in part I, contradicting the established theorem that $z^n = f(x, y)$ is regular for n a power of a prime. We deduce that $k(k+1)j \geqq m$; *for any μ satisfying $k(k+1)\mu < m$, curves of order $m - 3 - \mu$ with the points $O(k-1)$-fold are regular.*

To state the general case, let $m - 3 - j$ be the maximum order for which curves with the points $O(k - 1 - i)$-fold do not form a regular system. We find that

$$k(k+1)j \geqq (ik + i + 1)m.$$

The geometric result may be stated thus: *If $f(x, y) = 0$ be irreducible of order m, with singular points O, each of which is the origin of one branch of order k and class 1, and if μ be the maximum integer such that $k(k+1)\mu < (ik + i + 1)m$, then all complete systems of curves of orders $m - 3$, $m - 4, \cdots, m - 3 - \mu$ with the points $O(k - 1 - i)$-fold are regular.* Similar theorems can be stated for curves having the points $O(k - 1 - i)$-fold, and having in addition at each point O a given number τ of tangents coinciding with the tangent to f at that point, using in each case the value for d given in (19). In particular, if μ is the maximum integer satisfying $k(k+1)\mu < (k^2 - k - 1)m$, then curves of order $m - 3 - \mu$ passing simply through the points O form a regular system; if μ is the maximum integer given by $k(k+1)\mu < (k^2 - 2k - 2)m$, curves of order $m - 3 - \mu$ passing through the points O in contact with f at each point O form a regular system.

12. *Conditions on such singularities of f.* The theorem just stated

imposes certain conditions on the number and position of singularities O possible on an irreducible curve of given order. The question of existence of an irreducible curve possessing a stated set of singularities can in some cases be given an immediate negative answer, by this theorem, whereas the extended Plücker relations would be cumbersome to handle or would yield no conclusive answer. A simple example shows this: A curve of order 13 can have 22 ordinary triple points; when these become origins of a single branch each, the Plücker relations reduce this number to 19. Now by the theorem just proved, curves of order $m-3-\mu$ with these points as double points must form a regular system for $k(k+1)\mu < m$; but the system of curves of order 9 with 19 double points has virtual dimension less than -1, hence 19 such points are impossible for an irreducible curve of order 13. As an example of the second type of restriction, consider curves of order 10, with points O each the origin of a single branch of order 3 and class 1; cubics simply through these points must form a regular system. There cannot be eleven such points, then; there are at most ten and if there are nine, they cannot be base points of a pencil of cubics.

13. *One branch of class greater than* 1. The two cases treated in detail make clear the fact that if we know the inequalities governing exponents of terms occurring in the equation of the adjoint surface ϕ_ν, we can describe the systems of curves $|C^d|$ cut out in the plane π, and enunciate for these systems theorems on regularity for determined orders. When O is the origin of a branch of order k and class c prime to k, the governing inequality * obtained by a transformation of the integral is of the form

$$(22)\qquad kn(\alpha+\beta-k+2-c)+nc(\beta+1)+k(k+c)(\gamma+1)>0.$$

Under this condition the curves $|C^d|$ must have the following properties: The generic C^0 has points O as $(k-1)$-fold points; all tangents at the point are arbitrary if the class $c \leq k/k-1$; one tangent coincides with the tangent to f if $k/k-1 < c \leq k/k-2$, etc. Further, the contact may be more than ordinary for any branch of C^0 which has contact with f, when c and k satisfy relations readily obtained from (22). All systems given by

$$d=0,1,\cdots,\ [nc/k(k+c)]-1$$

have the points $O((k-1)$-fold; the next set, given by

* This inequality can be obtained directly from the description of the behavior of the adjoint surfaces at O as given by the Newton polyhedron for F, as well as by a transformation exactly analogous to that of paragraph 8: See W. V. D. Hodge, "The isolated singularities of an algebraic surface," *Proceedings of the London Mathematical Society*, ser. 2, Vol. 30, part 2 (1929).

$$d = [nc/k(k+c)], \cdots, [n(k+3c)/k(k+c)] - 1$$

have the points $O(k-2)$-fold, and so on. The systems $|C^0|, \cdots, |C^{n-2}|$ whose superabundances may be involved in expressing the irregularity of F do not necessarily include systems with points O as ordinary $(k-i)$-fold points with arbitrary tangents. The sequence contains a group of systems $|C^{\delta_1}|$ to $|C^{\delta}|$ having the points O as $(k-i)$-fold points and satisfying certain other conditions as to number of tangents coinciding with the tangent to f and as to the nature of the contact; these additional conditions decrease as we progress from δ_1 to δ, but they do not necessarily disappear entirely before the appearance of $|C^{\delta+1}|$ with $O(k-i-1)$-fold. However, if we denote by s_δ the superabundance of $|C^\delta|$, then a system $|C|$ of the same order having the points $O(k-i)$-fold with arbitrary tangents will have superabundance $s \leq s_\delta$, hence if $|C^\delta|$ must be regular, certainly $|C|$ is regular. In this way we are able to enunciate theorems analogous to those of part II on the regularity of curves of determined order with behavior at the points O specified.

III. Application to f having Points of different Multiplicities.

14. In I and II, the behavior of the curves $|C^d|$ is described at points O for any value of d. The theorems there are stated for the case of f with multiple points all of the same order, but when different multiplicities are present, it is necessary only to describe the behavior for separate multiplicities and then state the general theorem as the intersection of these theorems. Let $f(x,y) = 0$ be irreducible of order m, with ordinary multiple points $O_1, \cdots, O_r$ of orders $k_1, \cdots, k_r$ respectively. From paragraph 6 we deduce the theorem: *All complete systems of curves of order $m-3-\mu$ with the points O_i as $(k_i-2-\rho_i)$-fold points are regular provided*

$$k_i\mu < (\rho_i+1)m. \qquad (i = 0, \cdots, r) \quad (\rho_i = 0, 1, \cdots, k_i - 2).$$

Suppose f has, in addition, points O_h each the origin of a single branch of order k_h and class 1. Each order k_i determines a set of critical values of d

$$d = [n/k_i], \ [2n/k_i], \cdots, \ [(k_i-2)n/k_i], \qquad (i = 0, 1, \cdots, r)$$

for which the multiplicity of points O_i on C^d decreases by one; similarly, each k_h determines a set of values

$$d = [n/k_h(k_h+1)], \cdots, \ [(k_h^2 - k_h - 1)n/k_h(k_h+1)]$$

for which the multiplicity of points O_h on C^d decreases by one. If these values are arranged in order of magnitude, the sequence $|C^0|, |C^1|, \cdots$ may be

described successively, for each time that a critical value of d is reached, the subscript will show which points change their behavior on $|C^d|$ for that value of d. The systems with base points thus specified will be regular for order $m-3-\mu$, where μ can be determined by the method of paragraph 6. For example, suppose f has cusps and ordinary triple points. The critical values of d arranged in order of magnitude give the behavior of $|C^d|$ as follows:

d	0	$[n/6]$	$[n/3]$
Multiplicity of C^d at triple points	1	1	0
Multiplicity at cusps	1	0	0

Consequently, curves of order $m-3-\mu$ passing simply through the triple points and cusps are regular for $6\mu < m$; curves of order $m-3-\mu$ passing simply through the triple points only are regular for $3\mu < m$.

The nature of the curves $|C^d|$ occurring in the sequence for any given f depends on the relation of the multiplicities $k_1 \cdots k_r$ among themselves, since n may be taken as large as we please. If, for example, f has multiple points of orders k_1, k_2 respectively, where $k_1 = \rho k_2$, the multiplicity of $|C^d|$ at points O_1 will have reached $k_1 - 2 - \rho$ when the multiplicity at points O_2 decreases to $k_2 - 3$. The systems of curves obtained by this method are essentially different, then, from the curves obtainable by the process of successive adjunction to f.

IV. Curves with Non-cyclic Fundamental Groups.

15. *Sufficient conditions that the fundamental group of f be non-cyclic.* In his paper on cyclic multiple planes,* Zariski proved that if the fundamental group of a plane irreducible algebraic curve $f(x, y) = 0$, in general position with respect to the line at infinity, is cyclic, the surface $z^n = f(x, y)$ is regular for every positive n. From this and the expression of the irregularity in terms of the superabundances of the systems of curves through the cusps, he deduces a connection between the structure of the fundamental group and the position of the cusps of f when f has nodes and cusps only. In any case, a condition under which $z^n = f(x, y)$ is irregular will be a sufficient condition that f have a non-cyclic fundamental group.

When f has only ordinary k-fold points, the systems of curves $|C^d|$ fall into consecutive sets each with the points $O(k-i)$-fold; if any of these systems are superabundant, F will be irregular. From paragraph 6, we see that if $m-3-j$ is the maximum order for which the points O as $(k-i)$-fold base points give rise to a superabundant system, then

* O. Zariski, "On the irregularity, etc.," *loc. cit.*

(23) $k\lambda_i = (h+1)\{kj_i - (i-1)m\} - k - k[j_i m_0/m] + (i-1)m_1 + \sigma.$

If $\lambda_i \geq 0$ for a particular value of i, no non-vanishing superabundance will appear due to those $|C^d|$ with points $O(k-i)$-fold; when $\lambda_i \geq 0$ for all values of i from 2 to $k-1$, no non-vanishing superabundance can appear in (12), and F is regular. From (23) we obtain

$$kj_i \geq (i-1)m \qquad (i = 2, 3, \cdots, k-1).$$

If $kj_i > (i-1)m$ for a particular i, the corresponding $\lambda_i \geq 0$; if then $kj_i > (i-1)m$ for all the values of i given, F is regular for $n \geq m$. It is then of necessity regular for values $n < m$, for if F were irregular for a value $n = n_1$, it would be irregular for n any multiple of n_1. For F to be irregular, therefore, it is necessary that at least one of the quantities

$$m,\ 2m,\ 3m, \cdots;\ (k-2)m$$

be divisible by k.

Suppose that for a particular value of i between 2 and $k-1$, we have $kj_i = (i-1)m$. From (23) we see that if $(i-1)m_1$ is not divisible by k, $\lambda_i \geq 0$; we must have not only $(i-1)m$ but also $(i-1)m_1$, and hence $(i-1)n$, divisible by k. When this is true, $\sigma = 0$ and $\lambda_i < 0$; we may state the result: *If $f(x,y) = 0$ is an irreducible algebraic curve of order m, with only ordinary k-fold points, the necessary and sufficient condition that the surface $z^n = f(x,y)$ $(n \geq m)$ be irregular is that $(i-1)m$ and $(i-1)n$ be divisible by k for at least one value of $i = 2, 3, \cdots, k-1$, and that putting $kj_i = (i-1)m$, the system of curves of order $m-3-j_i$ with the k-fold points of f as $(k-i)$-fold points be superabundant.*

For f having points of the second kind discussed, of course, more values of d enter as critical. We may state a set of sufficient conditions: for example,—*A sufficient condition that f, irreducible of order m, with points O_1 ordinary k-fold and points O_2 each the origin of a single branch of order k and class 1, have a non-cyclic fundamental group is that m be a multiple of $k(k+1)$, and that writing $m = k(k+1)j$, the systems of curves of order $m-3-j$ with $O_1(k-2)$*-fold and *$O_2(k-1)$*-fold *be superabundant.* The condition that f with nodes and cusps only have a non-cyclic fundamental group, mentioned at the beginning of this paragraph, is a special case of the above theorem. The set of sufficient conditions concludes with: *A sufficient condition that f, irreducible of order m, with points O_1 ordinary multiple points of order k and points O_2 each the origin of a single branch of order k and class 1, have a non-cyclic fundamental group is that $(k^2-k-1)m$ be divisible by $k(k+1)$, and that for $(k^2-k-1)m = k(k+1)j$, the curves of order $m-3-j$ passing simply through the points O_2 form a superabundant system.*

SECOND NOTE ON THE CELESTIAL SPHERE.

By Frank Morley.

1. *The caustic of a correspondence.* In a note on the Celestial Sphere [this Journal, Vol. 54 (1932), p. 276], I attempted to show the convenience of regarding the space around one as bounded by a sphere Ω. The points of Ω are named by numbers $x, y \cdots$. An arc orthogonal to Ω (or if preferred a line of the hyperbolic space) is named by its end points x and y.

To call Ω the Cayley absolute would be to begin at the wrong end—to pass from projective geometry to inversive geometry. The aim is rather to infer the properties of the included space from the known nature of its boundary.

We suppose now a correspondence between the points x and y, and that this is differentiable, so that for $x + dx$ we have in general $y + dy$.

We have a double infinity or congruence of arcs x, y. And the problem is to determine their envelope. This is, in a Euclidean space (when Ω is a point), a part of Hamilton's theory of systems of rays.

It is convenient to regard the sphere Ω as a plane. An arc x, y is then a semi-circle, say above the plane. The arc x, y intersects a consecutive arc $x + dx$, $y + dy$ when the pairs of points are on a circle and are interlaced. The cross-ratio $dxdy/(x - y)^2$ is then positive. We have then, when $d\theta$ is real,

$$dxdy/(x-y)^2 = (d\theta)^2; \tag{1}$$

thus the condition of intersection is that this differential invariant under homographies is to be invariant also under antigraphies (that is not altered by writing $\bar{x}$ for x) and is further to be positive. Geometrically stated, the elements dx and dy make opposite angles with $x - y$.

To find the point of intersection, we first find the point z where the join of x and y meets the join of $x + dx$ and $y + dy$. Taking the circle on which they lie as the base-circle, and replacing x and y by turns t and τ we are to have

$$z + \bar{z}t\tau = t + \tau$$

and

$$\bar{z}(\tau dt + t d\tau) = dt + d\tau.$$

Let $dt/d\tau = \mu e^{i\alpha}$, where μ is positive (the magnification).

Then $\tau^2 dt/t^2 d\tau = \mu e^{ia}$

$$dt/d\tau = \pm \mu t/\tau.$$

For interlaced pairs, dt/t and $d\tau/\tau$ have the same sign, so that, as we have taken μ positive,

$$dt/d\tau = \mu t/\tau.$$

We have then

$$\bar{z}(1+\mu) = 1/t + \mu/\tau$$
$$z(1+\mu) = t + \mu\tau,$$

and therefore in general

$$z = (x + \mu y)/(1+\mu), \tag{2}$$

where

$$\mu = | \, dx/dy \, | .$$

The point z is thus the internal center of similitude for corresponding small circles around x and y.

On the normal to Ω at this point z, we take a distance ζ given by

$$\zeta^2 = (z - x)(\bar{y} - \bar{z}) \tag{3}$$

to obtain the intersection of the consecutive arcs. The envelope of the arcs is then given by (1), (2), and (3). For convenience let it be called the *caustic* of the correspondence.

2. *The caustic of a homography.* As a case in point, consider the homography

$$x = \kappa y.$$

Then $dx = \kappa dy$. Then from (1)

$$\frac{\kappa dy^2}{(\kappa - 1)^2 y^2} = (d\theta)^2,$$

where $d\theta$ is real. Hence $y = c \exp \lambda\theta$, where $\lambda = \nu\kappa - 1/\nu\kappa$, c a constant.

$$x = \kappa c \exp \lambda\theta$$

$$z = \frac{\kappa + \mu}{1 + \mu} \, c \exp \lambda\theta$$

and

$$\zeta^2 = \frac{\mu(1-\kappa)(1-\bar{\kappa})}{(1+\mu)^2} \, c\bar{c} \exp (\lambda + \bar{\lambda})\theta,$$

that is $\zeta^2 = c_1 z\bar{z}$, where c_1 is a constant. This is a right cone. The arcs

then all touch this cone; the direction of an arc at contact makes a constant angle with the generator.

For the general homography

$$\alpha xy + \beta x + \gamma y + \delta = 0,$$

it follows that the arcs touch an inverse of a cone, that is a Dupin cyclide. The double points of the cyclide are the fixed points of the homography.

3. *The caustic of an antigraphy.* As a second case, consider the antigraphy

$$\alpha x\bar{y} + \beta x + \gamma \bar{y} + \delta = 0,$$

When there are fixed points, say 0 and ∞, this is

$$x = \mu\bar{y}.$$

The equations (1), (2), (3) are unaltered from (1)

$$\frac{\mu dy d\bar{y}}{(x-y)^2} = (d\theta)^2.$$

Hence $x - y$ is a real.

Hence x, y, and z are reals. Thus in this case the only arcs which can intersect consecutive arcs are those erected on the axis of reals. We have

$$\zeta^2 = (1-\mu)^2 z^2/4\mu$$

where z is a real; and the arcs touch two half-lines in the vertical plane on the axis of reals, the lines being images as to Ω. For any antigraphy with fixed points, the envelope of arcs is accordingly two arcs which meet Ω at the fixed points.

When the antigraphy has interchanging points 0, ∞, it may be taken

$$x\bar{y} = t^2$$

where t is a given turn.

From (1) we have

$$(d\theta)^2 = -\frac{x}{\bar{y}}\,\frac{dy d\bar{y}}{(x-y)^2},$$

so that

$$\frac{x}{\bar{y}(x-y)^2} = \frac{\bar{x}}{y(\bar{x}-\bar{y})^2},$$

that is $xy/(x-y)^2$ is a real.

Hence x and therefore y must be taken on the base circle. The envelope of the arcs is then a circle parallel to Ω, lying on the base sphere.

For an antigraphy with interchanging points, the envelope of arcs is accordingly a circle, such that the spheres on it, which touch Ω, touch it at the interchanging points.

In general for an algebraic correspondence $f(x, y) = 0$, that is for a Riemann surface, we shall have a caustic *surface*. The equation (1) giving at the place (x, y), or its correspondent (y, x), two directions at right angles, gives a natural dividing of the surface into orthogonal curves.

But for a correspondence $f(x, \bar{y}) = 0$, we shall have a caustic *curve*. In particular this is the case when the equation is self-conjugate, that is, when it is the image-system of an algebraic curve.

4. *The euclidean case.* When the infinity Ω becomes a point, the Dupin cyclide becomes a right cylinder and the arcs are all lines which touch this cylinder at a given angle.

For antigraphies in this case, the circles which cut Ω, other than orthogonally, disappear. We are left with a circle of the euclidean space. The arcs are lines such that, if 0 be the center of the circle, 1 the radius, p any interior point, they are perpendicular to the stroke $0, p$ and make with the disc an angle whose cosine is p.

ON THE COMMON POINTS OF TWO PLANAR CUBICS AND OF TWO PLANAR CYCLIDES.

By Frank Morley and W. K. Morrill.

We consider a pencil of cubics in a plane,

$$C_0 + \lambda C_\infty = 0.$$

The nine base points $(1, 2, 3, \cdots, 9)$ of the pencil are said to be associated and any one is determined by the other eight (supposed given). The problem is: Given eight real points and a ruler to construct the associated ninth.

Any quartic Q on the base points may be written as

$$C_0 L_\infty - C_\infty L_0 = 0$$

where L_0, L_∞ are linear.

The cubic $C_0 + \lambda C_\infty = 0$ meets Q, outside of the base points, at three points on a line of the pencil $L_0 + \lambda L_\infty = 0$; and this pencil of lines meets the quartic at a point c, the coresidual point of the base points as to Q. In particular if the quartic is two conics, one, B, on 16789, and the other, B', on 2345, any cubic of the system meets B' at two points, B at one point; these three points are on a line which meets B again at a fixed coresidual point c. Now there exist an infinity of conics on 2345 and to each there corresponds a point c on 16789, and conversely. In particular to a pair of lines on 2345, say 23 and 45, corresponds a point c_1 which we may indicate by:

$$16789 \mid 23 \mid 45.$$

This point is determined by 16789 and 23. Any cubic on these seven points meets the line 23 again at x, the conic again at y; and the line xy meets the conic again at the point c_1. Therefore we may denote c_1 by:

$$16789 \mid 23.$$

To construct it we may take for the cubic on the seven points a conic 26789 and a line 13. This can be done in ten ways. The points x and y are then given by Pascal's theorem.

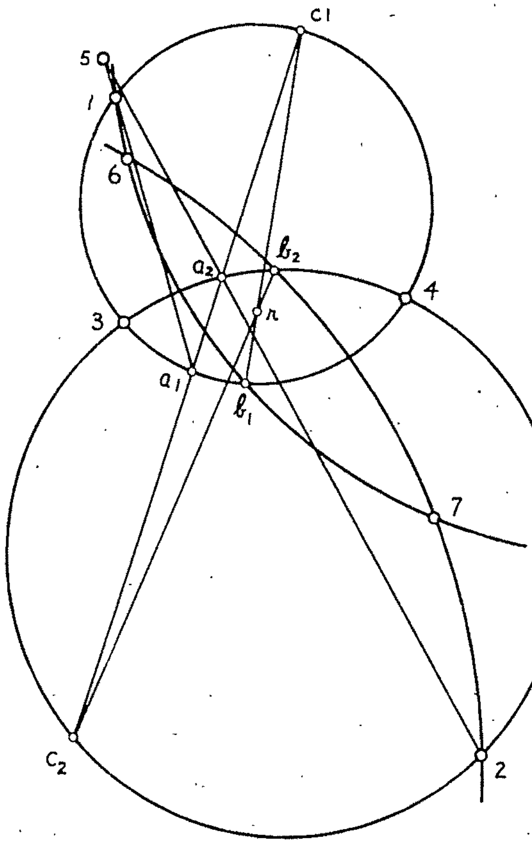
If 23 be the "circular points" and the conic 16789 be an ellipse, then using the eccentric angle θ the coresidual point is:

$$c_1 = \theta_1 + \theta_6 + \theta_7 + \theta_8 + \theta_9.$$

But the construction by Pascal's theorem fails.

The essential fact is that we have for the point c_1,

(1) $$16789 \mid 23 \equiv 16789 \mid 23 \mid 45 \equiv 16789 \mid 45.$$



Take now the two conics 16789 and 26789 and the point 5. We have the two degenerate cubics

$$16789 \times 25$$

and $$26789 \times 15$$

meeting again at two points a_1 and a_2. The join of these points meets 16789 at c_1 or $16789 \mid 25$ and meets 26789 at c_2 or $26789 \mid 15$. Let us suppose now that 5 is the sought point. We recall that $16789 \mid 25$ is identically $16789 \mid 34$, and $26789 \mid 15$ is identically $26789 \mid 34$. Thus we have the construction of 5: Find the points $16789 \mid 34$ and $26789 \mid 34$. Their join

determines the points a_1 and a_2 on the conics, and $a_1 1$, $a_2 2$ meet at the required point 5.

There is the related inversive problem of the eight intersections of two cyclides (or bicircular quartics). Here seven points determine mutually an eighth. If 8 and 9 in the projective case be taken as the "circular points" the cubics are "circular cubics" or in the inversive plane cyclides on the point ∞. Thus the above construction becomes inversive if we replace 8 and 9 by the single point ∞. We have now the eight associated points 1234567∞ and wish to construct the point 5, when the other seven are given.

We take the pairs of circles

$$\begin{cases} 134 \quad \text{and} \quad 234 \quad \text{on} \quad 34. \\ 167 \quad \text{and} \quad 267 \quad \text{on} \quad 67 \end{cases}$$

Let the two radical axes 34 and 67 meet at r. Let 167 and 134 meet again at b_1. Let 267 and 234 meet again at b_2. Let rb_1 meet 134 again at c_1, and rb_2 meet 234 again at c_2. Let the line c_1c_2 meet 134 at a_1 and 234 at a_2. Then $a_1 1$ and $a_2 2$ meet at 5.

It is to be noticed that in the inversive problem we do not have to assume actual intersections. A pair of circles may or may not intersect. Thus 3 and 4 may be a common image pair. And it would not be difficult to handle the case when there are no actual intersections, that is when there are four common image-pairs.

For other solutions one may consult P. Serret, *Géométrie de Direction*, H. S. White, *Cubic Curves*, and the references given in H. F. Baker's *Principles of Geometry*.

THE RECTANGULAR FIVE-POINT.

By F. Morley and R. C. Yates.

I. An Involution-Form.

We take four points on a plane as the points μ_i, $1/\mu_i$ of a rectangular hyperbola. They are on a circle if

$$\mu_1\mu_2\mu_3\mu_4 = 1$$

and are orthocentric if

$$\mu_1\mu_2\mu_3\mu_4 = -1.$$

The Euler line of $\mu_1\mu_2\mu_3$ is then in rectangular coördinates

$$\begin{vmatrix} x & y & 1 \\ -1/s_3 & -s_3 & 1 \\ s_1 & s_2/s_3 & 3 \end{vmatrix} = 0$$

and meets the curve at

$$\begin{vmatrix} \mu & 1/\mu & 1 \\ -1/s_3 & -s_3 & 1 \\ s_1 & s_2/s_3 & 3 \end{vmatrix} = 0$$

that is, at $\mu + 1/s_3 = 0$, the orthocenter, and at a point μ_0 given by

$$\begin{vmatrix} \mu_0 & s_3 & 0 \\ 1 & s_3^2 & -s_3 \\ s_1 & s_2/s_3 & 3 \end{vmatrix} = 0$$

or, explicitly

(1) $$\mu_0(\mu_1\mu_2 + \mu_2\mu_3 + \mu_3\mu_1 + 3\mu_1^2\mu_2^2\mu_3^3) = 3\mu_1\mu_2\mu_3 + (\mu_1 + \mu_2 + \mu_3)\mu_1^2\mu_2^2\mu_3^2.$$

Now this is for a given μ_0 an involution-form. That is, for given μ_1 and μ_2 we have three values of μ_3, say μ_3, μ_4, μ_5 which are symmetrically related to μ_1 and μ_2. For we have from (1)

$$\begin{aligned} \mu_1^2\mu_2^2(\mu_3 + \mu_4 + \mu_5) &= 3\mu_0\mu_1^2\mu_2^2 - \mu_1^2\mu_2^2(\mu_1 + \mu_2) \\ \mu_1^2\mu_2^2(\mu_3\mu_4 + \mu_4\mu_5 + \mu_5\mu_3) &= 3\mu_1\mu_2 - \mu_0(\mu_1 + \mu_2) \\ \mu_1^2\mu_2^2\mu_3\mu_4\mu_5 &= \mu_0\mu_1\mu_2 \end{aligned}$$

that is, writing s_i for a product-sum of all five,

(2) $$s_5 = \mu_0$$
(3) $$s_1 = 3s_5$$
(3′) $$s_4 = 3.$$

Five points on a rectangular hyperbola so related give ten triangles, as $\mu_1\mu_2\mu_3$; the Euler lines of the ten all meet at the point μ_0.

The equations (3) and (3′) are central in the paper of Mrs. Dean * so that we have here another way of approach to the theorems there given. It is there shown that the relations (3) and (3′) belong to the six points, μ_i and ∞. A six-point a_i is completed by drawing the six biquadratics which have a double point at one point and are on the other five. And in the case in question each curve cuts itself at right angles, or is rectangular. It is fair then to call this six-point *rectangular*, and to call a five-point, which with ∞ makes a rectangular six-point, also rectangular. Our five-point μ_i is then rectangular, in the sense that the 'circular cubics' on the five points and with a double point at one of them, will all cut themselves at right angles.

II. Neuberg's Invariant.

For four points, the condition that the four Euler lines meet at a point is found by eliminating μ_5 from (3) and (3′) to be

$$\sigma_1\sigma_3 - 10\sigma_4 + 3\sigma_4^2 + 3 = 0.$$

It is shown (p. 594 of the memoir cited) that this invariant is a factor of Neuberg's invariant for a four-point:

$$\begin{vmatrix} \lambda_{12}\lambda_{34} & \lambda_{12}+\lambda_{34} & 1 \\ \lambda_{23}\lambda_{14} & \lambda_{23}+\lambda_{14} & 1 \\ \lambda_{31}\lambda_{24} & \lambda_{31}+\lambda_{24} & 1 \end{vmatrix},$$

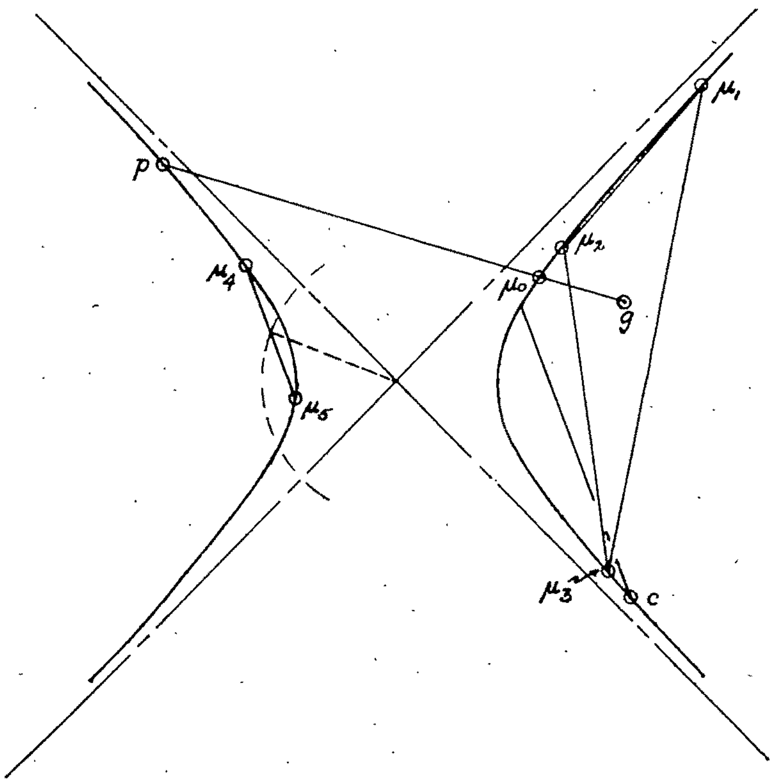
where λ_{ij} is the squared distance of two points. Accordingly, we may say that when the Euler lines all meet, for a four-point, the Neuberg invariant vanishes, or else the four-point is concyclic.

III. Construction of the Rectangular Five-Point.

It is of interest to construct a rectangular five-point. We take on a rectangular hyperbola three points μ_1, μ_2, μ_3 and mark the orthocenter p, the centroid g, and the point μ_0 where their join meets the curve again.

---

* *American Journal of Mathematics*, Vol. 52 (1930), p. 592.



We have from (2) and (3)

$$\mu_0 - (\mu_1 + \mu_2 + \mu_3)/3 = (\mu_4 + \mu_5)/3$$

and from (2) and (3′) the same in the reciprocals $1/\mu$. Hence, if we draw from the center of the hyperbola a stroke 3/2 that from g to μ_0 we have the mid-point of μ_4 and μ_5. And from (2) the chord μ_4, μ_5 is parallel to the chord μ_0, c where c is the point $1/\mu_1\mu_2\mu_3$.

THE QUARTIC SPACE INVOLUTORIAL TRANSFORMATIONS WITH A DOUBLE CONIC.*

By L. A. Dye and F. R. Sharpe.

1. *Introduction.* Some of the birational transformations of space which are determined by quartic surfaces with a double conic, are briefly developed by Cremona's method in a paper by Aroldi.† Among these transformations there is a $(4-4)$ transformation with a double conic, a rational quartic curve meeting the conic in four points, and a single isolated fundamental point. Miss Aroldi gives a short synthetic discussion of the fundamental elements and of the principal surfaces of the transformation but does not derive the analytical form nor does she consider the possibility of the transformation being involutorial.

In this paper the equations of the involutorial quartic transformation with a double conic and a rational quartic curve are found by showing that the transformation is the product of a quadratic transformation and a bilinear cubic transformation. If the quadratic transformation is involutorial and certain restrictions are imposed on the cubic transformation, then the quartic transformation is involutorial and its equations are obtained by this method.

It is also shown that there is another type of involutorial transformation with a double conic which is the transform of a linear involutorial transformation by a quadratic transformation.

2. *The Cremona transformation.* $F_4 : C_2{}^2 + C_4(p=0) + 0$. Two quartic surfaces F_4 in a three space S_3 having a double conic C_2 in common, meet in a residual curve of order eight, C_8. By the formulas of Noether ‡ this C_8 meets C_2 in eight points and is of genus five. If the two surfaces have a rational quartic curve L_4 in common, meeting C_2 in four points, the C_8 consists of L_4 and a C_4 which, by Noether's formulas, meets C_2 and L_4 in four and six points respectively and is of genus zero. Conversely through C_4 passes ∞^4 quartic surfaces having C_2 as a double conic, such that any two meet in an L_4 which meets C_4 in six points and C_2 in four points. Three of

* Presented to the Society, December 28, 1931.

† G. Aroldi, *Giornale di Matematiche di Battaglini*, ser. 3, Vol. 58 (1920), pp. 175-192.

‡ M. Noether, *Annali di Matematica*, ser. 2, Vol. 5 (1871), pp. 163-177.

these quartic surfaces meet in two points not on C_2 or C_4, so that if one point is fixed, there is a homaloidal web of surfaces, $F_4 : C_2^2 + C_4 + 0$. There exists therefore a Cremona transformation, T_4 between two spaces S_3 and S'_3 by which a plane of S'_3 corresponds to a surface of the web of F_4 in S_3.

If the web of $F_4 : C_2^2 + C_4 + 0$ of the transformation T_4 is subjected to a quadratic involutorial transformation I_2 having C_2 and 0 as fundamental elements, a web of cubic surfaces $F_3 : C_2 + C'_4(p = 0)$ is obtained. The composite curve $C_2 + C'_4$ is a $C_6(p = 3)$, and the web of cubic surfaces define a bilinear Cremona transformation T_3. Hence the quartic transformation may be obtained as the product of a bilinear T_3 and a I_2 having a conic in common. The variable curve of intersection of two cubic surfaces of the web of F_3 is a C_3 meeting C'_4 in six points and C_2 in two points. The ∞^1 trisecants of C'_4 are P-lines (P = principal) and they generate a quadric surface P_2 containing C'_4. The bisecants of C'_4 which meet C_2 are also P-lines, and their locus is a ruled surface of order six which contains C_2 as a triple conic and C'_4 as a double quartic, $P_6 : C_2^3 + C'^2_4$. These two P-surfaces make up the jacobian $J_8 : C_2^3 + C'^3_4$. Any cubic surface of the web meets P_2 in C'_4 and two P-lines, and meets P_6 in C_2^3, C'^2_4, and four P-lines. Hence P_2 corresponds to a conic C_2'', P_6 corresponds to a quartic curve C_4'', and the fundamental curve in S'_3 consists of the conic C_2'' and the rational quartic curve C_4'' which meets C_2'' in four points. The transformation T_3^{-1} is therefore similar to T_3.

Since $T_4 \equiv T_3 I_2$, it follows that if T_4 is involutorial $T_3 I_2 \equiv I_2 T_3^{-1}$, and hence that T_3^{-1} must have $C_2 + C_4$ as fundamental sextic, and that 0 must be invariant under T_3^{-1}. Amongst the web of F_3 there is a net of $F_3 : C_2 + C'_4 + 0$ which, by I_2, go into the plane of C_2 and a net of $F_3 : C_2 + C_4 + 0$. Amongst the web of F_2 there is a net of quadrics, each consisting of a plane through 0 and the plane of C_2, which are transformed by T_3^{-1} into a net of $F^3 : C_2 + C_4 + 0$ and the plane of C_2. Hence the net of planes through 0 are transformed by T_4 into the net of cubics $F_3 : C_2 + C_4 + 0$ and the plane of the conic. There is a composite cubic surface in the web of F_3 which consists of the plane of the conic and the unique quadric surface $P_2 : C'_4$. Under I_2 this goes into a unique quartic surface $P_4 : C_2^2 + C_4 + 0^2$. The plane of C_2 is transformed by I_2 into the cone $K_2 : C_2 + 0^2$ which goes by T_3^{-1} into the quartic surface $P_4 : C_2^2 + C_4 + 0^2$. Therefore by the involution T_4 the plane of the conic must go into this quartic surface.

3. *The web of* $F_3 : C_2 + C'_4$. The transformations T_3 and T_3^{-1} can be derived from three equations bilinear in (x) and (x'),

$$(1)\qquad \begin{aligned} u_1x'_1 + u_2x'_2 + u_3x'_3 + u_4x'_4 &= 0, \\ v_1x'_1 + v_2x'_2 + v_3x'_3 + v_4x'_4 &= 0, \\ w_1x'_1 + w_2x'_2 + w_3x'_3 + w_4x'_4 &= 0. \end{aligned}$$

When we put $x_4 = 0$, the four cubic curves, obtained by equating to zero the four third order determinants of the matrix

$$(2)\qquad \left\| \begin{matrix} u_1 & u_2 & u_3 & u_4 \\ v_1 & v_2 & v_3 & v_4 \\ w_1 & w_2 & w_3 & w_4 \end{matrix} \right\|,$$

must be of the form $(a_1x_1 + a_2x_2 + a_3x_3)K_2 = 0$, where $K_2 = 0$ is the cone $K_2 : C_2 + 0^2$. The two surfaces

$$\begin{aligned} u_1(v_3w_4 - v_4w_3) + v_1(u_4w_3 - u_3w_4) + w_1(u_3v_4 - u_4v_3) &= 0, \\ u_2(v_3w_4 - v_4w_3) + v_2(u_4w_3 - u_3w_4) + w_2(u_3v_4 - u_4v_3) &= 0, \end{aligned}$$

obtained from (1), determine a pencil of $F_3 : C_2 + C'_4$, passing through a residual C_3 which meets the conic $C_2 \equiv x_4 = K_2 = 0$ in two points P and Q and the plane $x_4 = 0$ in one other point R. If we put $x_4 = 0$ we get, by putting certain restrictions on u_1, u_2, v_1, v_2, w_1, and w_2, the conic C_2 through P and Q, and a pencil of lines through R. The matrix (2) is then of the required form. By linear combinations of rows and columns and by a linear transformation on (x), this matrix can be reduced to a normal form,

$$(3)\qquad \left\| \begin{matrix} 0 & -x_3 & x_2 & c_1x_1 \\ x_3 & 0 & -x_1 & c_2x_2 \\ -x_2 & x_1 & 0 & c_3x_3 \end{matrix} \right\|.$$

The four cubic equations obtained from it are $x_1K_2 = 0$, $x_2K_2 = 0$, $x_3K_2 = 0$, and a fourth which is identically equal to zero. The cone K_2 has the form $c_1x_1^2 + c_2x_2^2 + c_3x_3^2 = 0$. From (3) a general form for (1) can be obtained in the following manner; replace the first three columns in (3) by the three columns of the determinant which is the product of

$$\left| \begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{matrix} \right| \equiv D \quad \text{and} \quad \left| \begin{matrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{matrix} \right|,$$

and add terms in x_4 with coefficients from the elements of

$$\left| \begin{matrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{matrix} \right| \equiv \Delta.$$

The three bilinear equations now have the form

$$
\begin{aligned}
&(a_{13}x_2 - a_{12}x_3 + b_{11}x_4)x'_1 + (a_{23}x_2 - a_{22}x_3 + b_{12}x_4)x'_2 \\
&\qquad\qquad + (a_{33}x_2 - a_{32}x_3 + b_{13}x_4)x'_3 + c_1x_1x'_4 = 0, \\
&(-a_{13}x_1 + a_{11}x_3 + b_{21}x_4)x'_1 + (-a_{23}x_1 + a_{21}x_3 + b_{22}x_4)x'_2 \\
&\qquad\qquad + (-a_{33}x_1 + a_{31}x_3 + b_{23}x_4)x'_3 + c_2x_2x'_4 = 0, \\
&(a_{12}x_1 - a_{11}x_2 + b_{31}x_4)x'_1 + (a_{22}x_1 - a_{21}x_2 + b_{32}x_4)x'_2 \\
&\qquad\qquad + (a_{32}x_1 - a_{31}x_2 + b_{33}x_4)x'_3 + c_3x_3x'_4 = 0.
\end{aligned}
$$

The equations of T_3 as obtained from these bilinear equations are

$$
\begin{aligned}
T_3 :\ & x'_i = (A_{ij}x_j)K_2 + M_ix_4 + (c_jB_{ji}x_j)x_4^2, \qquad (i = 1, 2, 3), \\
& x'_4 = x_4[(b_{j1}x_j)(A_{1j}x_j) + (b_{j2}x_j)(A_{2j}x_j) + (b_{j3}x_j)(A_{3j}x_j) + Lx_4 + \Delta x_4^2],
\end{aligned}
$$

where $K_2 \equiv c_1x_1^2 + c_2x_2^2 + c_3x_3^2$,

$$
\begin{aligned}
M_i &\equiv \text{a quadratic ternary form in } x_1,\ x_2,\ x_3, \\
L &\equiv \text{a linear ternary form in } x_1,\ x_2,\ x_3, \\
(A_{ij}x_j) &= A_{i1}x_1 + A_{i2}x_2 + A_{i3}x_3, \\
(c_jB_{ji}x_j) &= c_1B_{1i}x_1 + c_2B_{2i}x_2 + c_3B_{3i}x_3, \text{ etc.}
\end{aligned}
$$

The A_{ij}, B_{ji} are cofactors of a_{ij}, b_{ij} in D and Δ respectively. The equations of T_3^{-1} are

$$
\begin{aligned}
T_3^{-1} :\ & x_i = \phi'_i = (a_{ji}x'_j)K'_2 + M'_ix'_4 + [c_1c_2c_3(b_{ij}x'_j)x'^2_4]/c_i, \quad (i = 1, 2, 3), \\
& x_4 = x'_4\phi'_4 = x'_4[c_1(a_{j1}x'_j)^2 + c_2(a_{j2}x_j)^2 + c_3(a_{j3}x_j)^2 + c_1c_2c_3x'^2_4],
\end{aligned}
$$

where $K'_2 \equiv (b_{1j}x'_j)(a_{j1}x'_j) + (b_{2j}x'_j)(a_{j2}x'_j) + (b_{3j}x'_j)(a_{j3}x'_j)$,

$$
\begin{aligned}
M'_i &\equiv \text{a quadratic ternary form in } x'_1,\ x'_2,\ x'_3, \\
(a_{ji}x'_j) &= a_{1i}x'_1 + a_{2i}x'_2 + a_{3i}x'_3, \\
(a_{j1}x'_j)^2 &= (a_{11}x'_1 + a_{21}x'_2 + a_{31}x'_3)^2, \text{ etc.}
\end{aligned}
$$

Let the involutorial quadratic transformation I_2 be

$$
\begin{aligned}
I_2 :\ & x'_i = x_ix_4, \qquad (i = 1, 2, 3), \\
& x'_4 = K_2 = c_1x_1^2 + c_2x_2^2 + c_3x_3^2.
\end{aligned}
$$

The quartic transformation obtained from the product T_3I_2 is

$$
\begin{aligned}
T_4 :\ & x'_i = x_4[(c_jB_{ji}x_j)K_2 + M_ix_4 + (A_{ij}x_j)x_4^2], \qquad (i = 1, 2, 3), \\
& x'_4 = \Delta K_2^2 + LK_2x_4 \\
&\qquad + [(b_{j1}x_j)(A_{1j}x_j) + (b_{j2}x_j)(A_{2j}x_j) + (b_{j3}x_j)(A_{3j}x_j)]\, x_4^2,
\end{aligned}
$$

and the inverse transformation T_4^{-1} obtained from $I_2T_3^{-1}$ is

$$T_4^{-1}: x_i = x'_4[(a_{ji}x'_j)K'_2 + M'_i x'_4 + \{c_1c_2c_3(b_{ij}x'_j)x'^2_4\}/c_i], \quad (i = 1, 2, 3),$$
$$x_4 = [c_1\phi'^2_1 + c_2\phi'^2_2 + c_3\phi'^2_3]/\phi'_4.$$

The transformations T_4 and T_4^{-1} must have the same form when T_4 is involutorial, hence

(4_1) $K_2 \equiv K'_2$,

(4_2) $a_{ji} = c_j B_{ji}$,

(4_3) $(c_1c_2c_3b_{ij})/c_i = A_{ij}$,

(4_4) $M_i \equiv M'_i$,

(4_5) $(c_1\phi'^2_1 + c_2\phi'^2_2 + c_3\phi'^2_3) \equiv [\Delta K_2^2 + LK_2x_4 + \{(b_{j1}x_j)(A_{1j}x_j) + (b_{j2}x_j)(A_{2j}x_j) + (b_{j3}x_j)(A_{3j}x_j)\}x_4^2]\phi'_4$.

If condition (4_2) is used to express the a_{ji} in terms of c_j and b_{ji}, then condition (4_3) requires that $\Delta = 1$. With these restrictions on a_{ij} and b_{ij}, condition (4_1) is satisfied if $b_{ij} = b_{ji}$, and no further restrictions are necessary in order that conditions (4_4) and (4_5) shall be satisfied. The quartic involutorial transformation now has the form

$$\begin{aligned} I_4: x'_1 &= x_4[(c_jB_{j1}x_j)K_2 + x_4\{c_2(c_jB_{j2}x_j)(b_{j3}x_j) - c_3(c_jB_{j3}x_j)(b_{j2}x_j)\} \\ &\quad + c_2c_3(b_{j1}x_j)x_4^2], \\ x'_2 &= x_4[(c_jB_{j2}x_j)K_2 + x_4\{c_3(c_jB_{j3}x_j)(b_{j1}x_j) - c_1(c_jB_{j1}x_j)(b_{j3}x_j)\} \\ &\quad + c_1c_3(b_{j2}x_j)x_4^2], \\ x'_3 &= x_4[(c_jB_{j3}x_j)K_2 + x_4\{c_1(c_jB_{j1}x_j)(b_{j2}x_j) - c_2(c_jB_{j2}x_j)(b_{j1}x_j)\} \\ &\quad + c_1c_3(b_{j3}x_j)x_4^2], \\ x'_4 &= K_2^2 + x_4^2[c_2c_3(b_{j1}x_j)^2 + c_1c_3(b_{j2}x_j)^2 + c_1c_2(b_{j3}x_j)^2], \end{aligned}$$

where B_{ij} are cofactors of b_{ij} in

$$\Delta \equiv \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} = 1, \qquad b_{ij} = b_{ji},$$

and $K_2 = c_1x_1^2 + c_2x_2^2 + c_3x_3^2$.

4. *The Cremona transformation* $F_4: C_2^2 + C_4(p = 0) + 0_1^2 + 0_2$. If the quartic curve C_4 has a double point 0_1^2 it can be shown that the quartic surfaces F_4 which contain it and a double conic C_2, also have the double point 0_1^2. The web of quartic surfaces F_4 may be transformed by a quadratic transformation $T_2: C_2 + 0_1$ into a web of quadric surfaces $F_2: C'_2 + 0_2$, hence $T_4T_2 \equiv S_2$ or $T_4 \equiv S_2T_2^{-1}$; that is, the quartic transformation may be regarded as the product of two quadratic transformations. Miss Hudson *

* H. Hudson, *American Journal of Mathematics*, Vol. 35 (1913), pp. 183-188.

has discussed this case and has derived the equations of the transformation.

If T_4 is involutorial, then $S_2T_2^{-1} \equiv T_2S_2^{-1}$ and S_2^{-1} has the same fundamental elements as T_2^{-1}. If S_2 is

$$
\begin{aligned}
(5) \qquad & x'_i = (a_{ij}x_j)(a_{4j}x_j), \qquad (i = 1, 2, 3;\ j = 1, 2, 3, 4),\\
& x'_4 = H_2[(a_{1j}x_j),\ (a_{2j}x_j),\ (a_{3j}x_j)],
\end{aligned}
$$

then under the transformation S_2^{-1}

$$
\begin{aligned}
(6) \qquad & (i_{ij}x_j) \sim x'_i x'_4, \qquad (i = 1, 2, 3;\ j = 1, 2, 3, 4),\\
& (a_{4j}x_j) \sim K_2(x'_1, x'_2, x'_3).
\end{aligned}
$$

Since $S_2S_2^{-1} = \text{Identity}$, we can substitute equations (6) in (5) and obtain

$$
\begin{aligned}
x'_i &= x'_i x'_4 K_2(x'_1, x'_2, x'_3), \qquad (i = 1, 2, 3),\\
x'_4 &= H_2(x'_1, x'_2, x'_3)x'^2_4,
\end{aligned}
$$

hence $H_2 \equiv K_2$.

The transformation S_2^{-1} is therefore

$$
S_2^{-1}:\ x_i = (A_{ji}x'_j)x'_4 + A_{4i}K_2, \qquad (i = 1, 2, 3, 4;\ j = 1, 2, 3),
$$

where the A_{ij} are cofactors of a_{ij} in the determinant

$$
\begin{vmatrix}
a_{11} & a_{12} & a_{13} & a_{14}\\
a_{21} & a_{22} & a_{23} & a_{24}\\
a_{31} & a_{32} & a_{33} & a_{34}\\
a_{41} & a_{42} & a_{43} & a_{44}
\end{vmatrix} \equiv A.
$$

The transformation S_2^{-1} is evidently the product of a linear transformation T_1 of determinant A, and the quadratic transformation T_2,

$$
\begin{aligned}
T_2:\ x'_i &= x_i x_4, \qquad (i = 1, 2, 3),\\
x'_4 &= K_2(x_1, x_2, x_3).
\end{aligned}
$$

Hence if T_4 is involutorial, $T_2T_1T_2^{-1} \equiv T_2T_1^{-1}T_2^{-1}$, and T_1 must be involutorial. The transform of a linear involutorial transformation by a quadratic transformation is a quartic involutorial transformation with a double conic and a quartic curve with a double point. It has the form

$$
\begin{aligned}
I_4:\ x'_i &= [(a_{ij}x_j)x_4 + a_{i4}K_2][(a_{4j}x_j)x_4 + a_{44}K_2], \quad (i = 1, 2, 3;\ j = 1, 2, 3, 4),\\
x'_4 &= (c_1a_{14}^2 + c_2a_{24}^2 + c_3a_{34}^2)K_2 + 2[c_1a_{14}(a_{1j}x_j) + c_2a_{24}(a_{2j}x_j)\\
&\quad + c_3a_{34}(a_{3j}x_j)]x_4K_2 + [c_1(a_{1j}x_j)^2 + c_2(a_{2j}x_j)^2 + c_3(a_{3j}x_j)^2]x_4^2,
\end{aligned}
$$

where the a_{ij} are elements of $A = |a_{ij}|$, $a_{ij} = A_{ji}$, and

$$
K_2 \equiv c_1x_1^2 + c_2x_2^2 + c_3x_3^2.
$$

A RELATION BETWEEN METRIC AND EUCLIDEAN SPACES.*

By W. A. Wilson.

1. *Introduction.* A semi-metric space is a set of points such that with each pair of points there is associated a unique non-negative number which is called the distance between them and is zero if, and only if, the points are identical. As Menger points out in his various articles on metric geometry,† a metric space is merely a semi-metric space which has the property that any three points are congruent with some three points in a Euclidean plane.

It is a natural inference from the above, as well as from Menger's general theorems, that a metric space in which any four points are congruent with some four points of a Euclidean space more nearly approximates Euclidean space than does a simple metric space. This inference is further strengthened by certain other results obtained by Menger (See II, pp. 209, 213). The present article is devoted to a further discussion of this question; the principal results are given in §§ 11 and 12.

2. *Notation.* A Euclidean space of n dimensions will be denoted by E_n. If a set A in a metric space Z is congruent with a set A' in a metric space Z', we say that A *can be imbedded in* Z'. If for a particular value of n any n points of the metric space Z can be imbedded in some Euclidean space, we say that Z has the *n-point property*.‡ Obviously, if n points can be imbedded in any Euclidean space, they can be imbedded in E_{n-1}.

The term *congruent* has its usual meaning and the fact that A is congruent with A' is denoted by the notation $A \simeq A'$. If $A \subset B$, $A' \subset B'$, $A \simeq A'$, $B \simeq B'$, and in every case the correspondence of a point a of A with a point a' of A' (Notation: $a \sim a'$) in the congruence $A \simeq A'$ is preserved in the congruence $B \simeq B'$, we say that the first congruence is *contained in* or is a *sub-congruence of* the second. If in the preceding definition the

* Presented to the Society, March 26, 1932.

† K. Menger, I. "Untersuchungen über allgemeine Metrik," *Mathematische Annalen*, Vol. 100, pp. 75-163, and Vol. 103, pp. 466-501; II. "Bericht über metrische Geometrie," *Jahresbericht der Deutschen Mathematiker-Vereinigung*, Vol. 40, pp. 201-219; III. "New Foundation of Euclidean Geometry," *American Journal of Mathematics*, Vol. 53 (1931), pp. 721-745. These will be referred to as I, II, or III. Many of the principles of Euclidean hyper-geometry used later are conveniently collected in III, pp. 727-729.

existence of the congruence $A \simeq A'$ implies the existence of one and only one congruence $B \simeq B'$ containing the given congruence, we say that the congruence $A \simeq A'$ *determines* the congruence $B \simeq B'$. If $\{a_i\}$ and $\{a'_i\}$ are two sets of points and we say that $\sum a_i \simeq \sum a'_i$, it will be understood that in this congruence $a_i \sim a'_i$ for each i. Similar remarks apply to congruences between two simple arcs ab and $a'b'$, two triangles abc and $a'b'c'$, two tetrahedra, and in general any two simplices.

A space Z is *complete* if any Cauchy sequence $\{a_i\}$ contained in Z converges to a unique point a in Z. If a and b are two points and c is a third point such that $ac + cb = ab$,* we say that c is *between* a and b. If for any two points of a metric space Z there is a third point between them, Z is called *convex*. If for any pair of points a and b of Z there are points c and d, such that b is between a and c and a is between b and d, Z is called *externally convex*.†

3. THEOREM. *In a metric space let $a_1, a_2, \cdots, a_n$ be n points which can be imbedded in E_{n-1}. If for some $k < n$ there are points $a'_1, a'_2, \cdots, a'_k$ in some E'_{n-1} such that $\sum_1^k a_i \simeq \sum_1^k a'_i$, then the n given points can be imbedded in E'_{n-1} so that the given congruence is preserved.*

Proof. By hypothesis there are n points $\{a_i''\}$ in E_{n-1} such that $\sum_1^n a_i \simeq \sum_1^n a_i''$. Also we have $\sum_1^k a_i'' \simeq \sum_1^k a'_i$. By a principle of Euclidean hypergeometry there is at least one congruence between E_{n-1} and E'_{n-1} which contains the congruence $\sum_1^k a_i'' \simeq \sum_1^k a'_i$. In the congruence $E_{n-1} \simeq E'_{n-1}$ let the images of the respective points $\{a_i''\}$, $k+1 \leqq i \leqq n$, be $\{a'_i\}$, $k+1 \leqq i \leqq n$. Then we have $\sum_1^n a'_i \simeq \sum_1^n a_i'' \simeq \sum_1^n a_i$, which was to be proved.

4. THEOREM. *In a metric space let $a_1, a_2, \cdots, a_n$ be n points which can be imbedded in E_{n-1}. If $k \leqq n$ and the set $\{a_i\}$, $1 \leqq i \leqq k$, can be imbedded in E_{r-2}, $r \leqq k$, then the set $\{a_i\}$, $1 \leqq i \leqq n$, can be imbedded in $E_{n-(k-r)-2}$.*

Proof. Let $\sum_1^n a_i \simeq \sum_1^n a'_i$, where the points $\{a'_i\}$ lie in an E_{n-1}. From

* Here, as often, ab denotes the distance between a and b. The use of ab also to denote a segment will cause no confusion.

† For general properties of convex spaces see Menger, I.

‡ This property is enjoyed for every value of n by any Euclidean space and by Hilbert space. See Menger, II, p. 215.

the second hypothesis we see that the set $\{a'_i\}$, $i \leqq k$, lies in an E_{r-2} contained in E_{n-1}. The space E_{r-2} and the points $\{a'_i\}$, $i > k$, which do not lie in E_{r-2} determine a Euclidean space of dimensionality at most $r-2+n-k$. Hence the theorem is proved.

Remark. Thus, if five points can be imbedded in E_4 and three of them can be imbedded in a line or four in a plane, all five can be imbedded in E_3; and if four of them can be imbedded in a line, all five can be imbedded in a plane.

5. *Properties of lines.* Let Z be a convex complete space which has the four-point property. If a and b are points of Z and $\{x\}$ is the set for which $ax + xb = ab$, it has been shown elsewhere that this set is congruent with a Euclidean segment of length ab. Similarly the set of points $\{x\}$ for which $ax + xb = ab$, or $ab + bx = ax$, or $xa + ab = xb$ is congruent to a Euclidean segment, ray, or open line and may be properly called a ***line in*** Z.

Other properties of lines in Z are: (a) any two points determine a line; (b) a line is determined by any two of its points; (c) three lines having one point in common can be imbedded in E_3; and (d) if a, b, and c are three points, the union of the segments ab, ac, and bc can be imbedded in E_2.

Finally, if Z is also externally convex, a line in Z is congruent with an open line in Euclidean space; i. e., it extends indefinitely in both directions.*

6. *Simplices in convex complete space.* Let Z be a convex complete space and $a_0, a_1, a_2, \cdots, a_k$ be $k+1$ points which can be imbedded in E_k but not in E_{k-1}. Let $a_1, a_2, \cdots, a_k$ be the vertices of a $(k-1)$-dimensional simplex S_{k-1} in Z. Let x be any point of S_{k-1}. Then the union of the segments $\{a_0x\}$ as x ranges over S_{k-1} is called a ***k-dimensional simplex in*** Z with vertices $a_0, a_1, a_2, \cdots, a_k$.

We must first show that this definition by recurrence has sense. We first note that by § 4 any r of the given points can be imbedded in E_{r-1} but not in E_{r-2}. A single point is called a 0-dimensional simplex and may be denoted by S_0. For $k=1$ and $S_0 = a_1$, a 1-dimensional simplex is the union of all the segments $\{a_0a_1\}$, of which one at least exists, since Z is convex and complete. For $k=2$ we can take the point a_0 and the one-dimensional simplex S_1 whose vertices are a_1 and a_2 and build at least one 2-dimensional simplex S_2. Continue in this way until we reach the simplex S_k.

---

* The definition of line is due to Menger and the above properties are readily deduced from Menger's theorems (See I, pp. 75-113) and the four-point property. For details see a paper by the author, "On Angles in Metric Spaces," which will be published in the *Bulletin of the American Mathematical Society.*

The next three sections are devoted to showing that in a convex complete space Z having the k-point property any n points $\{a_i\}$, $n \leqq k$, which cannot be imbedded in E_{n-2} have the following properties *provided that* $k \geqq 4$: (1) they are the vertices of a unique $(n-1)$-dimensional simplex S_{n-1} in Z; (2) if they are congruent with a set $\{a'_i\}$ in E_{n-1}, this congruence determines a unique congruence of S_{n-1} with the Euclidean n-dimensional simplex having the points $\{a'_i\}$ as vertices which contains the given congruence; (3) if Z is also externally convex, the n given points determine a unique sub-set H_{n-1} of Z such that $H_{n-1} \simeq E_{n-1}$ and this congruence is determined by the congruence in (2). We first note that the above statement is not true if $k = 3$. For, let Z be the surface of a sphere and a, b, and c be any three points not on the same great circle. If distances are geodetic distances on the spherical surface, Z is convex, complete, and metric. In general the three points are the vertices of a spherical triangle which conforms with the above definition of a two-dimensional simplex S_2, but it is not congruent with the plane triangle which has sides of lengths ab, ac, and bc. If two of the points are diametrically opposed, an application of the above definition gives absurd results.

For simplicity the words segment, triangle, and tetrahedron will be used for one-, two-, and three-dimensional simplices.

7. THEOREM. *Let Z be a convex complete space which has the four-point property. Let a, b, and c be three points which cannot be imbedded in E_1. Then Z contains exactly one triangle T whose vertices are a, b, and c, and, if a', b', and c' are points in any Euclidean space such that*

$$a + b + c \simeq a' + b' + c',$$

this congruence determines a congruence of T with the Euclidean triangle T', whose vertices are a', b', and c'.

Proof. By § 5 each pair of points are the ends of a unique segment. As the point x ranges over the segment bc let T be the union of the segments $\{ax\}$. Also, if x and y are distinct points of bc, $(ax)(ay) = a$. For otherwise a, x, and y would lie on a line by § 5 and, since $x + y \subset bc$, a would be on the line determined by b and c, contrary to hypothesis.

We now proceed to prove that the given congruence determines a congruence $T \simeq T'$. By § 5, $ab + bc + ac \simeq a'b' + b'c' + a'c'$. Let the points u and v of T lie on the segments ax and ay, respectively, where $x + y \subset bc$. In the congruence just given let $x \sim x'$ and $y \sim y'$. Then $ax = a'x'$ and $ay = a'y'$. On $a'x'$ there is just one point u' for which $au = a'u'$. Let $u \sim u'$. Likewise let $v \sim v'$. We must prove that $uv = u'v'$.

By § 5 we have $ab + by + ay \simeq a'b' + b'y' + a'y'$. For the same reason $ax + xy + ay \simeq a'x' + x'y' + a'y'$; then $av + vx + ax \simeq a'v' + v'x' + a'x'$; this gives $uv = u'v'$.* Hence $T \simeq T'$ and, by construction, this contains the given congruence. There is no other congruence $T \simeq T'$ containing the given congruence, for this would involve a congruence of T' with itself leaving the vertices invariant, but not itself an identical transformation.

This does not complete the proof. For we have yet to show that, if a triangle were built up with one of the other segments as a "base", it would be identical with T. To do this we prove that any set P of points of Z which contains a, b, and c and is congruent with T' in such a way that $a \simeq a'$, $b \simeq b'$, and $c \simeq c'$ is identical with T and that the congruence $P \simeq T'$ is the same as $T \simeq T'$. Let u' be any point of T', and $u \sim u'$ and $u'' \sim u'$ in the respective congruences $T \simeq T'$ and $P \simeq T'$. We must show that $u'' = u$.

Let $a'u'$ produced meet $b'c'$ in x'. In the above congruences let u, x and u'', x'' correspond to u', x', respectively. Now $bx = b'x'$, $cx = c'x'$, and $bx + cx = b'c' = bc$ by the first congruence; and $bx'' = b'x'$, $cx'' = c'x'$, and $bx'' + cx'' = b'c' = bc$ by the second. Hence $x'' = x$. In like manner $au = a'u'$, $ux = u'x'$, and $au'' + u''x = a'x' = ax$; whence $u'' = u$.

Thus there is exactly one triangle T with the vertices a, b, and c and one congruence $T \simeq T'$ determined by the congruence $a + b + c \simeq a' + b' + a'$.

8. *N-dimensional convex complete spaces.* Let Z be a convex complete space which has the four-point property; let $a_0, a_1, a_2, \cdots, a_n$ be $n + 1$ points which can be imbedded in E_n but not in E_{n-1}; let these points be the vertices of an n-dimensional simplex S in Z, which is unique and is congruent with an n-dimensional simplex S' in E_n with vertices $a'_0, a'_1, a'_2, \cdots, a'_n$ such that $\sum_0^n a_i \simeq \sum_0^n a'_i$; and let H be the set of points $\{x\}$ in Z such that each x is collinear with two points in S. Then H is called the *n-dimensional subspace of Z determined by $\{a_i\}$*; for brevity we call H an n-space.

If $n = 1$, an n-space is a line in Z. Properties of lines in Z were given in § 5; analogous results will now be obtained for n-spaces where $n > 1$.

THEOREM I. *Let Z be a convex complete space which has the four-point property and let H be an n-space in Z determined by the points $\{a_i\}$, $i = 0, 1, 2, \cdots, n$. If the set $\{a_i\}$ is congruent with a set $\{a'_i\}$ in E_n, the congruence $\sum_0^n a_i \simeq \sum_0^n a'_i$ determines a congruence of H with a sub-set of E_n.*

---

* For special positions of u and v, as on the same segment or on the parimeter, a similar, though, briefer, procedure gives the same result.

Proof. By the above definition there is an n-dimensional simplex S in Z whose vertices are $\{a_i\}$ and this is congruent with the n-dimensional simplex S' in E_n whose vertices are the points $\{a'_i\}$ in such a manner that $\sum_0^n a_i \simeq \sum_0^n a'_i$. If x is a point of H, let x be collinear with the points u and v of S and, to fix the ideas, let v lie between u and x. The congruence $S \simeq S'$ gives points u' and v' in S' corresponding to u and v. On $u'v'$ extended take the point x' so that v' is between u' and x' and $ux = u'x'$; then $vx = v'x'$. Let $H' = \{x'\}$. We now proceed to show that $H \simeq H'$ with $x \sim x'$.

Let w be any point of S different from u and v and let $w \sim w'$ by the congruence $S \simeq S'$. By the four-point property and § 3 there is a point x'' in E_n such that $u + v + w + x \simeq u' + v' + w' + x''$. Since $u'x'' = ux = u'x'$, $v'x'' = vx = v'x'$, and $u'x'' = u'v' + v'x''$, we have $x'' = x'$ and so $wx = w'x'$. Hence $S + x \simeq S' + x'$ for each x.

We now show that the correspondence of x' to x does not depend on the choice of u and v. Suppose that x were also collinear with r and s, where $r + s \subset S$; and, to fix the ideas, let x be between r and s. By the previous paragraph $rx = r'x'$ and $sx = s'x'$. As $rs = rx + sx$, this gives $r's' = r'x' + s'x'$; thus x' is on $r's'$ with $r'x' = rx$, just as it would have been chosen if we had started from the collinearity of r, s, and x.

Now let y be some other point of H collinear with points h and k of S and let k be between h and y. By the congruence $S + x \simeq S' + x'$ we know that $h + k + x \simeq h' + k' + x'$ and by the four-point property and § 3 there is a point y'' in E_n such that $h + k + x + y \simeq h' + k' + x' + y''$. Hence

$$h'y'' = hy = h'y', \quad k'y'' = ky = k'y', \quad \text{and} \quad h'y'' = h'k' + k'y''.$$

As $h'y' = h'k' + k'y'$, this gives $y'' = y'$. Then $x'y' = x'y'' = xy$. This gives both uniqueness of correspondence and congruence between H and H'.

There is no other congruence of H with H' containing the given congruence, for this would involve a congruence of H' with itself which would leave the $n + 1$ points $\{a'_i\}$ invariant but not be an identical transformation. This is impossible since the points $\{a'_i\}$ do not lie in an E_{n-1}.

THEOREM II. *Let Z be a convex complete space which has the four-point property and let H be an n-space in Z determined by the points $\{a_i\}$, $i = 0, 1, 2, \cdots, n$. Then H is convex, complete, and closed, and contains every point collinear with any two of its points.*

Proof. We first show that, if x and y lie in H and z is collinear with x and y, then z lies in H. By the definition of H and Theorem I there is a set of points $\{a'_i\}$, $i = 0, 1, 2, \cdots, n$, in E_n, which are congruent with $\{a_i\}$ and

are the vertices of an n-dimensional simplex S' in E_n, and the congruence $\sum_0^n a_i \simeq \sum_0^n a'_i$ determines a congruence of S with S' and of H with a sub-set H' of E_n.

Let u' be any inner point of S and $u \sim u'$. Then by the application of the four-point principle as in the proof of Theorem I, we can easily show that, if z' is the point on the line through x' and y' such that $xz = x'z'$ and $yz = y'z'$, then $u + x + y + z \simeq u' + x' + y' + z'$, and so $uz = u'z'$. If v' is a point on $u'z'$ so near to u' that v' is within S', there is a corresponding point v in S and, as in the case of u, $vz = v'z'$. The relations $uz = u'z'$, $uv = u'v'$, and $vz = v'z'$ show that z is collinear with u and v and so belongs to H. This also proves that H is convex.

In order to prove the rest of the theorem we first note that, since Z is metric, every convergent sequence is a Cauchy sequence and, since Z is complete, every Cauchy sequence converges to a point in Z. Hence, in order to prove that H is closed and complete, we need only to show that, if $\{x_i\}$ is a sequence of points in H which converges to a point x, then x is in H.

Let u' be an inner point of the simplex S', $u \sim u'$, and $x_i \sim x'_i$. Then the sequence $\{x'_i\}$ converges to a point x' in E_n, $u'x'_i = ux_i$, and consequently $ux = u'x'$. Let r be so small that the n-dimensional sphere of radius r and center u' lies in S'. On each segment $u'x'_i$ take the point v'_i such that $u'v'_i = r$ and let $v_i \sim v'_i$. Then each v_i lies on ux_i and in S. Obviously the sequence $\{v'_i\}$ converges to a point v' and, if $v \sim v'$, $v_i \rightarrow v$. We then have $uv = u'v'$. Also $v'_i x'_i \rightarrow v'x'$, $v_i x_i \rightarrow vx$, and consequently $vx = v'x'$. The relations $uv = u'v'$, $vx = v'x'$, $ux = u'x'$, and $u'x' = u'v' + v'x'$ show that $ux = uv + vx$; hence u, v, and x are collinear. This proves that x lies in H.

THEOREM III. *Let Z be a convex complete space which has the four-point property and let H be an n-space in Z determined by the points $\{a_i\}$, $i = 0, 1, 2, \cdots, n$. Let $\{b_i\}$, $i = 0, 1, 2, \cdots, n$, be points of H which determine an n-space K. Then $K = H$.*

Proof. By the definition of K the points $\{b_i\}$ are the vertices of an n-dimensional simplex S_n and any point x of K is collinear with two points of S_n. By Theorem II and the definition of a simplex $S_n \subset H$, since $\{b_i\} \subseteq H$. Applying Theorem II again, we see that x lies in H. Hence $K \subset H$ and in like manner $H \subset K$.

THEOREM IV. *Let the complete space Z be convex and externally convex, and have the four-point property. Let H be an n-space in Z determined by the points $\{a_i\}$, $i = 0, 1, 2, \cdots, n$, which are congruent with the set $\{a'_i\}$, $i = 0, 1, 2, \cdots, n$, in E_n. Then this congruence determines a congruence of H with E_n.*

Proof. By Theorem I there is a sub-set H' of E_n such that $H \simeq H'$. We have to prove that $H' = E_n$. If x' is any point of E_n, there are points u' and v' of the simplex S' whose vertices are $\{a'_i\}$ such that u', v', and x' are collinear. To fix the ideas, let v' be between u' and x'. Since $u' + v' \subset H'$, there are corresponding points u and v in H. As Z is externally convex, the line through u and v contains a point x such that $ux = u'x'$ and v is between u and x. Hence x lies in H and corresponds to a point x'' in H'. Now $u'x'' = ux = u'x'$ and x'' lies on $u'v'$ produced. Then $x'' = x'$. That is, $E_n \subset H'$. As $H' \subset E_n$, the theorem is proved.

Many theorems analogous to those of Euclidean geometry can now be deduced from the four theorems just proved and will occur to the reader. Of these mention should perhaps be made of the fact that, if $k < n$ and $k+1$ points $\{a_i\}$ of an n-space H determine the k-space K, then $K \subset H$ and the congruence of H with E_n (or a sub-set H' of E_n) implies the congruence of K with the E_k (or a sub-set of E_k) determined by the points of E_n congruent with $\{a_i\}$. Also, if Z is convex and complete and has the four-point property, an n-space H_n determined by the set $\{a_i\}$, $i = 0, 1, 2, \cdots, n$, contains every point x such that the set $x + \{a_i\}$ can be imbedded in E_n.

9. THEOREM. *Let Z be a convex complete space which has the four-point property. Let $\{a_i\}$, $i = 0, 1, 2, \cdots, n$, be $n+1$ points which can be imbedded in E_n but not in E_{n-1}. Then Z contains exactly one n-dimensional simplex S whose vertices are the points $\{a_i\}$ and, if $\{a'_i\}$, $i = 0, 1, 2, \cdots, n$, are points in any Euclidean space such that $\sum_0^n a_i \simeq \sum_0^n a'_i$, this congruence determines a congruence of S with the n-dimensional Euclidean simplex S' whose vertices are $\{a'_i\}$.*

Proof. This has been proved in § 7 for $n = 2$. Hence the theorem is true by induction if its validity for $n - 1$ implies its validity for n. Let us assume then that it is true for $n - 1$, $n \geqq 3$, and every smaller integer.

Then any n of the points, as $a_1, a_2, \cdots, a_n$, are the vertices of a unique $(n-1)$-dimensional simplex in Z, which we denote by T_0, the subscript indicating here and later that all the points $\{a_i\}$ *except* a_0 are vertices. As the point x ranges over T_0, let S be the union of the segments $\{a_0x\}$, which are unique by § 5. Also, if x and y are distinct points of T_0, $(a_0x)(a_0y) = a$. For otherwise a_0, x, and y would be collinear by § 5 and a_0 would lie in the $(n-1)$-space in Z determined by the points $\{a_i\}$, $i = 1, 2, \cdots, n$. Such a space exists by § 8, Theorem I, in consequence of our assumption of the validity of the theorem now being proved for $n - 1$. This, however, contradicts the hypothesis that $\{a_i\}$, $i = 0, 1, 2, \cdots, n$, cannot be imbedded in E_{n-1}.

Now the congruence $\sum_0^n a_i \simeq \sum_0^n a'_i$ implies the congruences

$$(1^0) \qquad T_0 \simeq T'_0;$$
$$(1^1) \qquad T_1 \simeq T'_1;$$
$$(1^2) \qquad T_2 \simeq T'_2;$$
$$\cdots\cdots\cdots$$
$$(1^n) \qquad T_n \simeq T'_n;$$

since our theorem is supposed to be true for $n-1$ and every smaller integer. Here, of course, each T'_i denotes the Euclidean simplex with vertices corresponding to those of T_i. It should also be noted here that, since these congruences are determined by the congruence $\sum_0^n a_i \simeq \sum_0^n a'_i$, if T_{ij} is the $(n-2)$-dimensional simplex common to T_i and T_j, the congruence $T_{ij} \simeq T'_{ij}$ is a sub-congruence of both congruences (1^i) and (1^j).

Let x and y be any points of S; they lie on unique segments a_0u and a_0v, where u and v are in T_0 by the definition of S. In congruence (1^0) let $u \sim u'$ and $v \sim v'$, and let the line through u' and v' meet two of the $(n-2)$-dimensional faces of T'_0, say T'_{0i} and T'_{0j}, in points r' and s'; let $r \sim r'$ and $s \sim s'$. (A similar discussion holds for the special cases where u' and v' lie on the same $(n-2)$-dimensional face of T'_0.) By congruences (1^0), (1^i), and (1^j) we have $rs = r's'$, $a_0r = a'_0r'$, and $a_0s = a'_0s'$. Hence by § 7

$$(2) \qquad \text{triangle } (a_0, r, s) \simeq \text{triangle } (a'_0, r', s').$$

We already have a one to one correspondence between the segments $\{a_0u\}$ in S and the segments $\{a'_0u'\}$ in S'. By (2), $a_0u = a'_0u'$ and $a_0v = a'_0v'$. Hence there are points x' and y' on a'_0u' and a'_0v', respectively, for which $a_0x = a'_0x'$ and $a_0y = a'_0y'$. If we let $x \sim x'$ and $y \sim y'$, we have a one to one correspondence between S and S'. But by congruence (2), $xy = x'y'$. Hence we have proved that $S \simeq S'$. There is no other congruence $S \simeq S'$ containing the given congruence, for this would involve a congruence of S' with itself leaving the vertices invariant but not itself an identical transformation.

Now let P be any set in Z which contains the set $\{a_i\}$ and is congruent with S' in such a way that $a_i \sim a'_i$, $i = 0, 1, 2, \cdots, n$. Let x' be any point of S' and $x \sim x'$, $x'' \sim x'$ in the respective congruences $S \simeq S'$, $P \simeq S'$. If we can show that $x'' = x$, then $P = S$.

Let a'_0x' produced meet T'_0 in u'_1; $a'_1u'_1$ produced meet T'_{01} in u'_2; $a'_2u'_2$ produced meet T'_{012} in u'_3; etc., until we finally reach a point u'_{n-1} on the segment $a'_{n-1}a'_n$. (We may of course reach a point on an edge of S' earlier; if so, the following discussion is merely shortened.)

In the congruence $S \simeq S'$, let $x \sim x'$ and $u_i \sim u'_i$; in the congruence

$P \simeq S'$ let $x'' \sim x'$ and $u_i'' \sim u'_i$. Now $a_{n-1}u_{n-1} = a'_{n-1}u'_{n-1}$, $u_{n-1}a_n = u'_{n-1}a'_n$, and $a_{n-1}u_{n-1} + u_{n-1}a_n = a'_{n-1}a'_n$; while $a_{n-1}u''_{n-1} = a'_{n-1}u'_{n-1}$, $u''_{n-1}a_n = u'_{n-1}a'_n$, and $a_{n-1}u''_{n-1} + u''_{n-1}a_n = a'_{n-1}a'_n = a_{n-1}a_n$. Hence $u_{n-1} = u''_{n-1}$ and $a_{n-2}u_{n-1} = a_{n-2}u''_{n-1} = a'_{n-2}u'_{n-1}$. Since u'_{n-2} lies on $a'_{n-2}u'_{n-1}$, we can repeat the reasoning just employed and show that $u_{n-2} = u''_{n-2}$. Eventually we have $u_1 = u''_1$ and can then show that $x'' = x$, whence $P = S$.

Thus there is but one simplex with the vertices $\{a_i\}$, $i = 0, 1, 2, \cdots, n$; in other words the result is independent of the choice of "base" from the $(n-1)$-dimensional simplices $T_0, T_1, T_2, \cdots, T_n$.

This completes the proof. It is obvious that the above theorem is merely a generalization of Menger's theorem that any two points in a convex complete space are the ends of a segment. The four-point condition insures uniqueness, but as yet we need the assumption of an $(n+1)$-point condition to insure the existence of *any* simplex. We now turn to a discussion of this point.

10. Let $\{a_i\}$, $i = 0, 1, 2, \cdots, n$, be any $n+1$ points in a metric space and suppose that these points cannot be imbedded in E_{n-1}. Also, for any $k \leqq n$, suppose that any k of these points can be imbedded in E_{k-1}, but cannot be imbedded in E_{k-2}. It is shown by Menger's Theorems 7 and 8 (see I, pp. 132-133) that, if D_n is the determinant

$$\begin{vmatrix} 0 & 1 & 1 & . & 1 \\ 1 & 0 & (a_0a_1)^2 & . & (a_0a_n)^2 \\ 1 & (a_1a_0)^2 & 0 & . & (a_1a_n)^2 \\ . & . & . & . & . \\ 1 & (a_na_0)^2 & (a_na_1)^2 & . & 0 \end{vmatrix};$$

then the $n+1$ points can be imbedded in E_n if and only if sign $D_n \neq$ sign $(-1)^n$. (If $D_n = 0$, they can be imbedded in E_{n-1}.) If for any k of the points, $k \leqq n$, determinants $\{D_{k-1}\}$ similar to D_n are formed, no $D_{k-1} = 0$ and in each case sign $D_{k-1} \neq$ sign $(-1)^{k-1}$.

If the values of a_ia_j are fixed except for a_0a_1 and we set $x = a_0a_1$, the determinant when expanded has the form $Ax^4 + Bx^2 + C$, where A, B, and C are constants. Therefore a small change in the value of x does not alter the sign of D_n; also, if the change is small enough, none of the determinants $\{D_{k-1}\}$, $k \leqq n$, will be altered in sign. Hence, if $n+1$ points satisfy the conditions laid down at the beginning of this section, a change in the distance between any pair, if sufficiently small, will leave the validity of the conditions unaffected.

By expanding D_n we see that $A = -D_{n-2}$, where D_{n-2} is the determinant formed from the points $\{a_i\}$, $i \geqq 2$. By Menger's theorems* $A \neq 0$ and sign A

* Menger also gives an extensive discussion of the function $Ax^4 + Bx^2 + C$, which includes the following properties. See I, pp. 120-135, and II, pp. 737-743.

$= -\operatorname{sign} D_{n-2} = \operatorname{sign}(-1)^{n-2} = \operatorname{sign}(-1)^n$. If we allow $x = a_0a_1$ to vary, $D_n = Ax^4 + Bx^2 + C$ will vary and will have the sign of A for x large enough. Hence $\operatorname{sign} D_n = \operatorname{sign}(-1)^n$ for x large enough. From the form of the function D_n we see that there are at most two non-negative real values α and β of x for which $D_n = 0$. In the next section it is shown that these always exist. Then, if $\alpha < \beta$, it is readily seen from the nature of the function $Ax^4 + Bx^2 + C$ that $\operatorname{sign} D_n \neq \operatorname{sign}(-1)^n$ when $\alpha < x < \beta$.

Hence, if we can prove that the validity of the n-point condition implies that $\alpha < x < \beta$, it follows that the $n + 1$ points can be imbedded in E_n.

11. Theorem. *Let Z be a convex complete space which is also externally convex and has the four-point property. Then Z has the n-point property for every integer n.*

Proof. It is sufficient to show that, if Z has the k-point property for every $k \leqq n$, where n is a fixed integer greater than or equal to 4, then Z has the $(n+1)$-point property.

Let $a_0, a_1, a_2, \cdots, a_n$ be any $n + 1$ points of Z, $n \geqq 4$. If $k \leqq n$, any k of these points can be imbedded in E_{k-1}. If the other points are all collinear with a_0 and a_1, all can be imbedded in E_1. In the opposite case let them be numbered so that a_2 is not collinear with a_0 and a_1. Then by § 7, a_0, a_1, and a_2 are the vertices of a triangle S_2 and by § 8 they determine a plane H_2 in Z. If all the other points lie in H_2, all the points can be imbedded in E_2. In the opposite case let a_3 not be in H_2. Then no three of the points a_0, a_1, a_2, and a_3 are collinear by § 8, Theorem II, they are the vertices of a tetrahedron S_3 in Z by § 9, and they determine a 3-space H_3 in Z by § 8. If all the other points lie in H_3, all the points can be imbedded in E_3, and the theorem is proved because $n \geqq 4$. In the opposite case let a_4 not be in H_3. Then no four of the points a_0, a_1, a_2, a_3, and a_4 lie in the same plane; for, if a_4 were in the same plane as any three of the others, it would be in H_3, since H_3 contains the plane in Z determined by any three of its points. If $n = 4$, we then have the result that no n of the points can be imbedded in E_{n-2}, while any n can be imbedded in E_{n-1}.

If $n > 4$, the points a_0, a_1, a_2, a_3, and a_4 can be imbedded in E_4 by hypothesis, they are the vertices of a 4-dimensional simplex S_4, and they determine a 4-space H_4; etc. Eventually we reach a stage where all the points are imbedded in E_{n-1} or the points $a_0, a_1, a_2, \cdots, a_{n-1}$ determine a unique $(n-1)$-space H_{n-1} in Z which does not contain a_n. In the latter case no k of the points, $k \leqq n$, can be imbedded in E_{k-2}, while any n of them are the vertices of a unique $(n-1)$-dimensional simplex in Z.

Now $a_1, a_2, a_3, \cdots, a_n$ are the vertices of a unique $(n-1)$-dimensional simplex T_0 and $a_0, a_2, a_3, \cdots, a_n$ are the vertices of a unique $(n-1)$-dimensional simplex T_1. These have as a common face the $(n-2)$-dimensional simplex T_{01}.

Let E_{n-2} be an $(n-2)$-space in E_n containing points $a'_2, a'_3, \cdots, a'_n$ congruent with $a_2, a_3, \cdots, a_n$, and let T'_{01} be the Euclidean $(n-2)$-dimensional simplex congruent with T_{01} and having the vertices $a'_2, a'_3, \cdots, a'_n$. Since Z has the n-point property, it follows by § 3 that in any E_{n-1} containing E_{n-2} and contained in E_n there is a point a'_1 such that $\sum_1^n a_i \simeq \sum_1^n a'_i$ and $T_0 \simeq T'_0$, where T'_0 is the Euclidean $(n-1)$-dimensional simplex of vertices $\{a'_i\}$, $i = 1, 2, \cdots, n$. Likewise there is a point a'_0 such that

$$a_0 + \sum_2^n a_i \simeq a'_0 + \sum_2^n a'_i \quad \text{and} \quad T_1 \simeq T'_1.$$

Since a reflection about E_{n-2} is a congruence, each of the points a'_0 and a'_1 can be chosen in two ways. If a'_0 and a'_1 are on opposite sides of E_{n-2}, the segment $a'_0a'_1$ cuts E_{n-2} in a point b'. By § 8, Theorem IV, the $(n-1)$-space H_{n-1} determined by the vertices of T_0 is congruent with E_{n-1} by a congruence determined by that between the vertices of T_0 and T'_0. This congruence also determines the congruence $T_{01} \simeq T'_{01}$ and a congruence between the $(n-2)$-space H_{n-2} determined by the vertices of T_{01} and E_{n-2}. Let b be the point of H_{n-2} such that $b \sim b'$. Likewise the $(n-1)$-space K_{n-1} determined by the vertices of T_1 is congruent with E_{n-1}, $T_1 \simeq T'_1$, $T_{01} \simeq T'_{01}$, and $H_{n-2} \simeq E_{n-2}$. Hence in this congruence, too, $b \sim b'$. Since a'_0, b', and a'_1 are collinear, $a'_0a'_1 = a'_0b' + b'a'_1$. Also $a_0b = a'_0b'$ and $ba_1 = b'a'_1$ by the respective congruences $K_{n-1} \simeq E_{n-1}$ and $H_{n-1} \simeq E_{n-1}$. Hence the triangle axiom of metric geometry gives $a_0a_1 \leqq a_0b + ba_1 = a'_0a'_1$. Set $a'_0a'_1 = \beta$.

If a'_0 and a'_1 were so taken as to be on the same side of E_{n-2}, the discussion in the second paragraph of § 10 shows that there is no loss of generality in assuming that a'_0 and a'_1 are not equidistant from E_{n-2}. Then $a'_0a'_1$ produced will meet E_{n-2} in a point c', to which corresponds a point c in H_{n-2}, and in the same manner as in the previous paragraph we find that $a_0a_1 \geqq | a_0c - a_1c | = a'_0a'_1$. In this case set $a'_0a'_1 = \alpha$.

It is easy to show that $\alpha < \beta$. If $a_0a_1 = \alpha$ or $a_0a_1 = \beta$, the $n+1$ points $\{a_i\}$ can be imbedded in E_{n-1} and by Menger's Theorem 7 (see I, p. 132), the determinant D_n discussed in § 10 equals zero. We have seen in the two preceding paragraphs that $\alpha \leqq a_0a_1 \leqq \beta$. Hence by the results of § 10 the $(n+1)$-point condition is satisfied. This completes the proof.

12. Theorem. *Let Z be a convex complete separable space which is also*

externally convex and has the four-point property. Then Z is congruent with some E_n or with Hilbert space.

Proof. In consequence of § 11 the space has the n-point property for every value of n. Hence the determinant D_n (see § 10) for any $n + 1$ points has a sign different from that of $(-1)^n$ or is zero. If for every value of n there is some set of $n + 1$ points for which $D_n \neq 0$, it then follows by a theorem of Menger * that Z is congruent to Hilbert space.

In the contrary event there is a smallest integer n such that, for every set of $n + 2$ points, the determinant $D_{n+1} = 0$. Then every set of $n + 2$ points can be imbedded in E_n (see I, p. 132). Consequently Z can be imbedded in E_n (see I, p. 128). But for some set of $n + 1$ points the determinant $D_n \neq 0$ and so this set of points can be imbedded in E_n but not in E_{n-1}. Hence by § 8 they determine an n-space H_n congruent to E_n. As Z can be imbedded in E_n, $Z = H_n$ and the theorem is proved for this case.

13. Several questions are immediately suggested by the result just obtained. As remarked in § 1 a semi-metric space is metric if the 3-point condition is satisfied, *i. e.*, if the determinant D_2 formed from any three points is not positive. This happens to be the case if the triangle axiom is valid. In view of our last theorem it would be especially interesting to have an axiom similar to the triangle axiom and equivalent to the condition that the 4-point determinant $D_3 \geqq 0$.

The overwhelming effect of the 4-point condition lends new force to Menger's identification of metric space with semi-metric space satisfying the 3-point condition,—it shows how drastic the triangle axiom really is despite its simplicity and suggests the possibility that the proper point of departure in the study of abstract spaces is semi-metric rather than metric space.

It seems to the writer reasonable to suppose that the conditions given in the theorems of §§ 11 and 12 are stronger than needed. The results of this article originated in an attempt to find a convex space which satisfied the 4-point condition and was *not* Euclidean. It will be noted that, while the removal of the condition of external convexity in § 12 makes this theorem false, it seems not unlikely that there is still congruence with a convex part of some Euclidean or Hilbert space. The condition is of more serious importance in the demonstration of § 11, where it is needed to establish the existence of the point c in Z corresponding to c' in the next to the last paragraph. It may well be that the conclusion of the theorem can be established by some other method without using the condition of external convexity.

YALE UNIVERSITY.

---

* *Wiener Akademischer Anzeiger*, 1928, No. 12. See also III, p. 745.

ON THE CONSTRUCTION OF SIMPLE ARCS.

By Gordon T. Whyburn.

1. In this paper it is proposed, in the first place, to give a simple, complete, and essentially new demonstration for the well-known theorem * that every compact locally connected continuum is arcwise connected. All previous proofs for this fundamental proposition with which the author is familiar are alike in that the arc is constructed by means of a monotone decreasing sequence of simple chains of open sets or regions. Due to the fact that two successive links in such chains necessarily overlap in more than one point, thereby leaving open the possibility that the next chain in the sequence may oscillate widely within a given chain, a certain element of elusive seems unavoidable when these chains are used to construct arcs; and in to completely eliminate all such possibilities of oscillation and make the argument entirely convincing, some tedious details are necessary.

In the proof to be given below these difficulties are avoided by making use, in the construction, of chains of continua, called *regular chains,* defined as follows: C is a ***simple regular chain*** of continua joining two points a and b if C is the sum of a finite number of compact continua $a \subset X_1, X_2, \cdots, X_n \supset b$ such that any two successive links have exactly one common point, links that are not successive have nothing in common, and only the first and last links contain a and b respectively. If $\delta(X_i) < \epsilon$ for each i, we shall call C a simple regular ϵ-chain. We show first that any two points of a locally connected continuum are joinable by simple regular ϵ-chains of *locally connected continua* for any $\epsilon > 0$. Thus by simple recursion we can set up a monotone decreasing sequence of these chains from a to b such that the norm, ϵ, approaches 0, and their product is easily proved to be an arc from a to b. Clearly such regular chains represent better approximations to arcs than do the simple chains of open sets, because since each intersection point of two links in the chain separates the chain between a and b, every such point in any chain in the sequence must belong to the product set and thus to the final arc. No difficulties of oscillation can arise in this construction because, there being only one common point for any two successive links of a given chain, any later

* See R. L. Moore, *Transactions of the American Mathematical Society,* Vol. 17 (1916), p. 136; S. Mazurkiewicz, *Fundamenta Mathematicae,* Vol. 1 (1920), p. 201.

chain must pass through all of these intersection points in the same order as the links occur from a to b.

In the remaining sections of the paper we discuss other methods of constructing arcs and indicate ways of obtaining the more general arcwise connectivity theorem for complete locally connected spaces. In particular it is shown that the latter result may be obtained from a more general junction property of general locally connected spaces which can be demonstrated by means of regular chains.

2. We proceed now with our proof for the

ARCWISE CONNECTIVITY THEOREM. *Every two points a and b of a compact locally connected continuum M can be joined in M by a simple continuous arc.*

Proof. (1). For any $\epsilon > 0$, $M = \sum_1^n M_i$, where each M_i is a locally connected compact continuum of diameter $< \epsilon$.

For it is well known that M is the image under a uniformly continuous transformation T of the unit interval I. Thus if for the given ϵ, I is divided into a finite number of subintervals $I_1, I_2, \cdots, I_n$, each of diameter less than the corresponding δ_ϵ, then $T(I_1) = M_1$, $T(I_2) = M_2, \cdots, T(I_n) = M_n$ have the required properties.

(2). For any $\epsilon > 0$, any point x of M is an interior point (rel. M) of some compact locally connected subcontinuum $K_\epsilon(x)$ of M of diameter $< \epsilon$.

For, set $M = \sum_1^n M_i$ as in (1), where $\delta(M_i) < \epsilon/2$, and let $K_\epsilon(x)$ be the sum of all these continua M_i which contain x.

(3). Any two points a and b of a region * R in M can, for each $\epsilon > 0$, be joined in R by a simple regular ϵ-chain of compact locally connected continua.

For let S be the set of all points of R which can be so joined to a. Then S is open in R; for if x is any point of S, if C is a chain of the desired type in R from a to x, if X is the link of C containing x, and if $\sigma > 0$ is a number $< \epsilon - \delta(X)$ and also $< \rho[x, F(R) + C - X]$, then the chain obtained from C by replacing the link X by the set $X + K_\sigma(x)$ is an ϵ-chain of the desired type in R having x as an interior point. But S is also closed in R. For let y be any limit point of S in R. Let us choose $\sigma < \epsilon$ so that $K_\sigma(y) \subset R$. As y is an interior point of $K_\sigma(y)$ and thus $K_\sigma(y) \cdot S \neq 0$, hence there exists a

* By a *region* in M is meant any connected open subset of M.

simple regular ϵ-chain C of locally connected continua $a \subset X_1, X_2, \cdots, X_n \supset x$ in R from a to some point x of $K_\sigma(y)$ such that $X_i \cdot K_\sigma(y) = 0$ for $i < n$. Now if C contains y, then $y \subset S$. If not, then set $K_1 = K_\sigma(y)$, choose a number $\epsilon_1 < 1/2$ and also $< \rho(y, C)$ and, applying (1), express K_1 as the sum of a finite collection G_1 of locally connected ϵ_1-continua. Let H_1 be the sum of all the continua of G_1 which contain no point of X_n and let Y_1 be the component of H_1 containing y. Then G_1 contains at least one element, say K_2, which contains at least one point of Y_1 and at least one point of X_n. In general, for any $k > 1$, choose a positive number $\epsilon_k < 1/2^k$ and also $< \rho(Y_{k-1}, X_n)$ and express K_k as the sum of a finite collection G_k of locally connected ϵ_k-continua. Let H_k be the sum of all those continua of G_k which contain no point of X_n and let Y_k be the component of the set $Y_{k-1} + H_k$ which contains y. Then G_k contains at least one element, say K_{k+1}, which contains at least one point of each of the sets Y_k and X_n. Let $Y = \sum_1^\infty Y_k$. Then since $Y \cdot X_n = 0$, since each Y_k is closed, and $Y - Y_k \subset K_k$, it is seen at once that $\bar{Y} - Y = \prod_1^\infty K_k = p$, a single point; and thus Y is a compact locally connected ϵ-continuum having just the point p in common with X_n. Thus $[X_1, \cdots, X_n, \bar{Y}]$ is a simple regular ϵ-chain of compact locally connected continua in R from a to y. Hence every limit point y of S in R belongs to S and thus S is closed in R. As S is both open and closed in R and R is connected, we have $S = R$. Thus $b \subset S$ and (3) is proved.

We proceed now to construct an arc in M from a to b. By (3), a and b can be joined in M by a simple regular 1-chain C_1 of compact locally connected continua. Likewise for each $n > 1$, it follows by (3), since C_{n-1} is a locally connected continuum, that a and b can be joined in C_{n-1} by a simple regular $1/n$-chain of locally connected continua. Let $C = \prod_1^\infty C_n$. Then C is a simple arc from a to b, as will now be demonstrated. In the first place, since each C_n is a compact continuum containing $a + b$ and $C_n \supset C_{n+1}$, it follows that C is a compact continuum containing $a + b$. Furthermore, C is locally connected. To prove this, let x be any point of C and, for any n, let A_n be the sum of the (at most two) links of C_n which contain x and let B_n be the sum of those not containing x. Then since $A_n \cdot B_n$ consists of 0, 1, or 2 points according as A_n contains 2, 1, or 0 points of the set $a + b$, it follows that in any case $A_n \cdot C$ is connected; and since $\delta(A_n \cdot C) < 2/n$, it follows that C is locally connected. Finally, each point x of $C - (a + b)$ separates * a and

* That is, $C - x = C_a + C_b$, where C_a and C_b are mutually separated and contain a and b respectively.

b in C. For let P_n be the set of all points common to at least two links of C_n and let $P = \sum_1^\infty P_n$. Then clearly $P \subset C$ and every point of P separates a and b in C. Let x be any point of $C - P - a - b$. Then if x does not separate a and b in C, some component R of $C - x$ contains $a + b$. But by (3), R contains a compact continuum K containing $a + b$, and hence $K \supset P$; but this is impossible, because x is a limit point of P but not of K. Therefore x separates a and b in C and, by the Sierpinski definition, C is a simple arc from a to b.

3. *Remark.* Although in the above proof we showed that the set C is a simple arc from a to b by proving that it satisfies the Sierpinski definition of an arc, it is interesting to note that the method of construction of C lends itself readily to yield a direct proof that C is homeomorphic with the unit interval. For clearly the chain C_1 can, by choosing a sufficiently small norm and by combining the proper number of links, be so chosen that the set P_1 of all points common to at least two links of C_1 contains exactly $2^{v_1} - 1$ points, where v_1 is some integer. Similarly the chain C_2 may be so chosen that each of its links lies wholly in some link of C_1 and each link of C_1 contains exactly 2^{v_2} links of C_2, where v_2 is some integer uniform for all links of C_1, and thus so that the set P_2 of all points common to at least two links of C_2 contains exactly $2^{v_1+v_2} - 1$ points. We can continue in this way so that the set P_n of all points common to at least two links of C_n contains exactly $2^{v_1+v_2+\cdots+v_n} - 1$ points, and such that each link of C_{n-1} contains exactly 2^{v_n} links of C_n. Now if for each n we place the points of $a + P_n + b$ into correspondence with the numbers $m/2^{v_1+v_2+\cdots+v_n}$, $(m = 0, 1, 2, \cdots, 2^{v_1+v_2+\cdots+v_n})$, in the order in which they occur in C_n from a to b, we obtain a $(1-1)$ correspondence T between the set $\sum_1^\infty P_n$ and the set D of rational numbers of the form $m/2^n$. It is readily shown that both T and its inverse are uniformly continuous, and hence both may be extended to their limit points, thus giving a homeomorphism between C and the interval $(0, 1)$.

4. *Another method.* We note also that the arcwise connectivity theorem may be obtained in the following manner. First prove the theorem for the case of continua which are regular in the sense of Menger-Urysohn. This case results at once from the fact that any two points in a compact continuum can be joined in that continuum by an irreducible subcontinuum, because any subcontinuum of a regular curve is regular and hence locally connected, and any locally connected continuum irreducible between two points is an arc

joining these two points.* (The same argument proves the theorem for the case of hereditarily locally connected continua, *i. e.*, continua containing only locally connected subcontinua.) This case established, the general theorem is obtained by showing that any two points in a locally connected continuum lie together in a regular subcontinuum. This is done by constructing the set C as in § 2. That C is regular is self evident.

5. *Extension to complete spaces.* We turn now to the well-known † extension of the arcwise connectivity theorem to complete spaces. This is embodied in the following proposition which we shall call the

GENERALIZED ARCWISE CONNECTIVITY THEOREM. *Every two points a and b of a connected and locally connected complete space S can be joined in S by a simple continuous arc.*

In the opinion of the author this more general theorem can best be proved by reducing the general case back to the case of compact spaces M treated in § 2 by means of the following proposition: ‡

(A). *If K is any self-compact subset of a connected and locally connected metric space N, then there exists a subset H of N which is dense in K and which is the image under a uniformly continuous transformation of the set of all dyadic rational numbers on the unit interval.*

As has been pointed out by the author ‡ this theorem (A) yields at once

* A proof for this fact may be found in Hausdorff's *Mengenlehre* (1927), p. 222. It does not seem to be generally recognized that in this characterization of simple arcs no assumption whatever of compactness is necessary. In other words, *a metric space C is a simple arc from a to b if and only if (β) C is connected and contains a and b but no closed and connected proper subset of C contains both a and b and (ϵ) C is locally connected.* For if K denotes the set of all points in a set C satisfying (β) and (ϵ) which separate a and b in C, then it is known [See G. T. Whyburn, *Bulletin of the American Mathematical Society*, Vol. 33 (1927), p. 685; and *Transactions of the American Mathematical Society*, Vol. 32 (1930), p. 927] that $K + a + b$ is closed and compact. But we must have $K + a + b = C$; for if not, then some point x of $C - (a + b)$ fails to separate a and b, and hence by (β) cannot separate C at all; but then using (ϵ) and the well known simple chain lemma (see R. L. Moore, *loc. cit.*), we can obtain a connected chain E in C from a to b neither containing x nor having x as a limit point; and this contradicts (β), since E is then a proper subset of C.

Thus we see that the condition (α) that C be compact in the characterization [α, β, ϵ] of simple arcs given by Hausdorff (*loc. cit.*) is entirely superfluous.

† See R. L. Moore, *Bulletin of the American Mathematical Society*, abstract, Vol. 33 (1927), p. 141; K. Menger, *Monatshefte Für Mathematik und Physik*, Vol. 36 (1930), p. 210; Kuratowski, *Fundamenta Mathematicae*, Vol. 15, p. 301; and Aronszajn, *ibid.*, p. 228.

‡ See G. T. Whyburn, *American Journal of Mathematics*, Vol. 53 (1931), p. 753.

the fact that, in case N is a complete space, $\bar{H}$ is a compact locally connected continuum. Thus every two points of the space S in the generalized arcwise connectivity theorem lie together in a compact locally connected continuum M in S, and from the arcwise connectivity of M we deduce the arcwise connectivity of S.

This method of proving the arcwise connectivity theorem has additional advantages accruing from the great generality of proposition (A). In particular, (A) yields at once the celebrated theorem of Hahn-Mazurkiewicz to the effect that any compact locally connected continuum is the continuous image of an interval. Thus in a treatment of locally connected spaces the logical sequence of the results discussed above seems to be as follows: First prove theorem (A); from this we obtain as an immediate consequence the continuous image theorem of Hahn-Mazurkiewicz; next establish the arcwise connectivity theorem for compact spaces M; then with aid of (A) we obtain the generalized arcwise connectivity theorem.

6. *General Junction Theorem.* In conclusion we shall consider a very general proposition embodying a junction property of locally connected metric spaces in general from which all the above arcwise connectivity theorems follow as immediate consequences. If X is any metric space, for convenience we shall denote by X_0 the space obtained by *completing* X by adding * on new points corresponding to every fundamental sequence in X.

JUNCTION THEOREM. *Every two points a and b of a connected and locally connected metric space N can be joined in N by a set H which is the image under a $(1-1)$ and doubly † uniformly continuous transformation of the set D of dyadic rational numbers ‡ on $(0,1)$. Thus H_0 is a simple arc from a to b.*

This theorem is readily established with the aid of regular chains. With the aid of a theorem of the author's § it can be shown that any two points of such a space N can, for each $\epsilon > 0$, be joined in N by a simple regular ϵ-chain of locally connected sets (links) such that two successive links have just one point in common and any two links that are not successive are at a positive minimum distance apart. This established, we can proceed to construct the set $H = \sum_1^\infty P_n$ by exactly the same method as was outlined above in § 3.

* See Hausdorff, *Grundzüge der Mengenlehre*, (1914), p. 315.

† That is, both the transformation and its inverse are uniformly continuous.

‡ i.e., numbers of the form $m/2^n$.

§ See *American Journal of Mathematics*, Vol. 53 (1931), p. 439.

Clearly the set H_0 obtained from H by *completing* it is homeomorphic with the interval $(0, 1)$ and hence is an arc in N_c from a to b. Thus we have shown that *any two points a and b of any connected and locally connected metric space N can be joined in N_0 by a simple arc $H_0 = ab$ such that the points H of N belonging to ab are everywhere dense on ab.* It is obvious that in case N is complete so that we have $N_c = N$, or is homeomorphic with a complete space,* we obtain the generalized arcwise connectivity theorem (see § 5). In addition to this, however, we see by virtue of the Junction Theorem that even in the general locally connected spaces, without any assumptions as to completeness or compactness, we have all the machinery necessary for the construction of simple arcs. Indeed, we can construct sets H in these spaces which, when completed, become simple arcs, and these arcs will lie in the original spaces provided these spaces are assumed to be complete.

THE JOHNS HOPKINS UNIVERSITY.

* This is known to include all G_δ subsets of any complete space. See Alexandroff, *Comptes Rendus*, Vol. 178 (1924), p. 185; also Hausdorff, *Fundamenta Mathematicae*, Vol. 5, p. 146.

BICONNECTED AND RELATED SETS.

By P. M. Swingle.

In Section 1 of this paper a number of simple theorems concerning biconnected * and related sets are obtained. In Section 2 various types of continua are studied with respect to the property of being the sum of mutually exclusive punctiform connected subsets. For example it is seen that a bounded indecomposable continuum can be the sum of any integral number of such sets greater than one. In the last section euclidean spaces are studied as the sum of biconnected subsets. It is shown that a necessary and sufficient condition that a euclidean space be n dimensional is that it be the sum of $n+1$, but not of n, biconnected subsets.

1. *Simple theorems concerning biconnected sets.* In this section a number of problems concerning biconnected sets are considered and a number of simple theorems proved concerning them and related sets.

It readily is seen that in order that a connected set C contain a finite dispersion set H, it is necessary and sufficient that every connected subset of C contain a finite dispersion set. Consider now the case where each of these dispersion sets contains the same number of points.

Lemma 1. *There does not exist a connected set C, every proper connected subset of which contains the finite set H of n points, unless $n=1$.*

For if there exists such a set C, then every proper connected subset of C contains any one point h of H. Hence C is biconnected and has the dispersion point h. As a biconnected set contains at most one dispersion point,† $n=1$.

Lemma 2. *If every proper connected subset of a connected set C contains a primitive dispersion set of n points, where n is any given positive*

* A biconnected set is a connected set which is not the sum of two mutually exclusive connected subsets, where it is understood that a connected set contains more than one point. These sets were defined and studied first by B. Knaster and C. Kuratowski in their interesting paper, "Sur les ensembles connexes," *Fundamenta Mathematicae,* Vol. 2, pp. 206-253. The problem of whether a biconnected set must contain a dispersion point was proposed by C. Kuratowski, *Fundamenta Mathematicae,* Vol. 3, p. 322 (19). The results of the present paper were obtained partly in consideration of this problem.

† J. R. Kline, "A Theorem Concerning Connected Sets," *Fundamenta Mathematicae,* Vol. 3, pp. 238-239.

integer, then each of these proper connected subsets contains the same primitive dispersion set H, which is also a primitive dispersion set of C. Thus n must equal one.

Let H be the primitive dispersion set of C if such exists; otherwise let $H = 0$. If $H \neq 0$, it is seen that it must contain the n points of the primitive dispersion set Q of any connected subset K of C.* And similarly Q must contain $K \times H$. Also it is seen that, if C contains a finite dispersion set D, then D contains $H \neq 0$.†

(1) Consider the case where there exists a point x of $C - H$ such that $C - x$ is connected, and so has a primitive dispersion set Q of n points. Then the dispersion set $Q + x$ contains $H \neq 0$ and so $H = Q$.

(2) Consider now the case where for each point x of $C - H$, $C - x = C_1 + C_2$ separate.‡ As $C_1 + x$ is connected H consists of $2n$ points.

Suppose that for a point h of H, $C - h = C_1 + C_2$ separate. Then H consists of $2n - 1 = 2n$ points, which is a contradiction. Hence $C - h$ is connected, and so H consists of $n + 1$ points.

Therefore $n + 1 = 2n$, and so $n = 1$. Thus H consists of the two points h_1 and h_2 say. Then for $C - x = C_1 + C_2$ separate each C_i must contain one and only one point of H. Say C_i contains h_i $(i = 1, 2)$. Let $w \neq h_1$ be a point of C_1. Then $(C_1 + x) - w = C_{11} + C_{12}$ separate, as $C - w = W_1 + W_2$ separate for each $W_i + w$ must contain points of C_1, the one which does not contain x being contained in C_1 and the other, W_2 say, containing points of C_1 as $W_2 \times C_2$ does not have w as a limit point. But as $C_{1i} + w$ is connected and only one of the primitive dispersion set can be contained in h_1, and so in H, a contradiction is obtained. Hence this case does not exist.

Therefore every connected subset of C must contain the primitive dispersion set H of C, and so by lemma 1, H consists of one point.

The truth of the following theorem and its corollary is now evident.

THEOREM 1. *In order that a connected set C contain a dispersion point it is necessary and sufficient that every proper connected subset of C contain a dispersion point.*

---

* R. L. Wilder, "On the Dispersion Sets of Connected Point Sets," *Fundamenta Mathematicae*, Vol. 6, Theorem 4, p. 216.

† R. L. Wilder, *loc. cit.*, Theorem 1, p. 214.

‡ By this notation, $C - x = C_1 + C_2$ *separate*, is meant that $C - x$ is the sum of the two non-vacuous, mutually exclusive, subsets C_1 and C_2, neither one of which contains a limit point of the other.

COROLLARY 1. *If every proper connected subset of a connected set C contain a dispersion point, then C is biconnected.*

COROLLARY 2. *There does not exist a connected set C such that, for every point x of C, the set $C - x$ is connected and has a dispersion point.*

For let K be any proper connected subset of C and let y be a point of $C - K$. Then $C - y$ contains K. Let q be the dispersion point of $C - y$. Then $(C - y) - q$ is totally disconnected and so does not contain K, which then must have q as its dispersion point. Hence as every proper connected subset of C has a dispersion point, by the above theorem, C does also, which is a contradiction.

Two theorems will be proved now concerning the dispersion point of a connected subset of a biconnected set.

THEOREM 2. *If B is biconnected and q is a dispersion point of a connected subset, K, of B, then B is the sum of two biconnected subsets, having only q in common, each of which has q as a dispersion point.**

It is evident that this theorem is true, if $B - q$ is disconnected. Consider then the case where $B - q$ is connected. Let $K - q = K_1 + K_2$ separate. Let C be the maximal connected subset of $B - K_1$, which contains the connected set $K_2 + q$. Then $B - C$ is totally disconnected as B is biconnected. Then, if $(B - C) + q = Y_1 + Y_2$ separate, where Y_1 contains the connected set $K_1 + q$, $B - C = (Y_1 - q) + Y_2$ separate. Hence $C + Y_2$ is a connected subset of $B - K_1$ and so C is not a maximal connected subset of this set. Thus $B - C + q$ must be connected. Therefore B is the sum of the two biconnected subsets C and $B - C + q$, having only q in common, which is a dispersion point of each.

THEOREM 3. *If B is biconnected and q is a dispersion point of a connected subset K of B, where K is connected im kleinen at q, then no connected subset of B has a dispersion point other than q.*†

Assume that C is a connected subset of B, having dispersion point $p \neq q$. By theorem 2, B is the sum of two biconnected subsets, mutually exclusive except for p, where each of these has p as a dispersion point. Let E be the one of these biconnected subsets which contains q. Let $E - p = E_1 + E_2$

* Whether q must be in this case a dispersion point of B itself is an unsolved problem of interest in connection with the unsolved problem, suggested by Kuratowski, of whether a biconnected set must always contain a dispersion point.

† Whether the theorem is true if K is not connected im kleinen is an unsolved problem.

separate where E_1 contains q. As K is connected im kleinen at q, K contains a connected subset N, containing q, such that $N \times (E_2 + p) = 0$. But as $E_2 + p$ is also a connected subset of B, this is impossible. Therefore $p = q$.

It readily is seen that if B is a biconnected set which is disconnected by a finite subset H, then H contains a dispersion point of B.* It follows that a biconnected set cannot contain two distinct cut points. Thus it is true that the closure † B, of a biconnected set B, contains at most one cut point, which must be a point of B. Furthermore, if this closure is a continuous curve in a euclidean space, then every point of $\bar{B}$ is contained in a simple closed curve of $\bar{B}$.‡ However it is not true necessarily that $\bar{B}$ is cyclicly connected.

It is known that there exist continuous curves, every subcontinuum of which is a continuous curve, which contain punctiform connected subsets.§ It follows from the following theorem that such continuous curves are not the sum of a countable set of arcs.

THEOREM 4. *If p is a point of a punctiform connected set P, in a euclidean space, and R is a region (sphere), containing p, then $R \times \bar{P}$ is not contained in a countable set of simple continuous arcs.*

Suppose that $R \times \bar{P}$ is contained in the countable set of arcs $A_1, A_2, \cdots$, where A_1 contains p. If x is a point of $A_i \times P \times R$, then x must be a limit point of points of P not contained in A_i, for P otherwise would have to contain a subcontinuum of A_i. Hence p is a limit point of points of $(A_2 + A_3 + \cdots) \times P \times R$, which are not contained in A_1. Consider for example the case where x_1 is a point of this set which is contained in A_5, but in no A_j where j is less than 5. Then there exists a region R_1 containing x_1, where R contains R_1, such that $R_1 \times (A_1 + A_2 + A_3 + A_4) = 0$. In a similar manner there exists a point x_2 having the same relation to x_1 that x_1 has to p, where x_2 is contained in a region R_2, such that R_1 contains $\bar{R}_2$, which contains no point common with any arc with subscript less than j where j is now greater than 5. Similarly points $x_3, x_4, \cdots$ are obtained with corresponding regions $R_3, R_4, \cdots$ where each $\bar{R}_{i+1}$ is contained in R_i and has points common with fewer arcs than R_i did. Hence it is seen that the set $x_1, x_2, \cdots$ has a limit point in $\bar{P}$ which none of the arcs $A_1, A_2, \cdots$ contains.

---

* R. L. Wilder, *loc. cit.*, Theorems 1, 4, and 10.

† By the closure of a set B is meant the set $\bar{B}$ consisting of B and the limit point of B.

‡ This follows from Theorem 11 of W. L. Ayres' paper "Concerning Continuous Curves in Metric Space," *American Journal of Mathematics*, Vol. 51 (1929), p. 591.

§ B. Knaster and C. Kuratowski, "A Connected and Connected im Kleinen Point Set Which Contains No Perfect Subset," *Bulletin of the American Mathematical Society*, Vol. 33 (1927), pp. 106-109.

2. *Continua as the sum of punctiform connected subsets.* The following lemma can be proved readily by means of transfinite induction and Zermelo's postulate,* as it is possible to continue to choose points for the K_f's from the C_i's and for the Q_g's from the P_i's, since the product of the power of the set of points chosen and $m+2$ is less than n, except where this power is itself n.

LEMMA A. *Let h be the first ordinal number of power m and b be the first ordinal number of power n; where $mn=n$ and $m<n$. Let (C_i) and (P_i) $(i=1,2,\cdots b)$ be two classes of point sets, each C_i and each P_i being of power n. Then (1) there exists a class (K_f) $(f=1,2,\cdots,h)$, of power n, of mutually exclusive point sets K_f such that each K_f contains a point of every C_i of (C_i) and (2) there exists a class (Q_g) $(g=1,2)$ of mutually exclusive point sets Q_g such that each Q_g contains a point of every P_i of (P_i); also for each K_f and Q_g, $K_f\times Q_g=0$.*

THEOREM A. *Let W be a bounded indecomposable continuum, lying in a euclidean space, and let m be any cardinal number greater than one and less than c, the power of the linear continuum. Then W is the sum of m mutually exclusive, punctiform, connected subsets, each of which is dense in W.*†

For let (C) be the class of the continua of our space, each of which separate points of W, and let (P) be the class of the perfect sets of W. Both (C) and (P) have power c. Hence by Zermelo's well-ordering theorem ‡ these sets can each be well-ordered obtaining classes (C_i) and (P_i) $(i=1,2,\cdots,k)$, where the power of k is c.

Let a and b be any two points of W separated by C_i of (C_i) and let T be any *composant* § of W. Then as $T+a+b$ is connected, C_i must contain

---

* W. Sierpinski, *Leçons sur les Nombres Transfinis* (1928), pp. 164-165, p. 212, and p. 137.

† The method of proof used here is due to B. Knaster and C. Kuratowski, *Fundamenta Mathematicae*, Vol. 2, pp. 250-251. It is to be noted that if any of the above punctiform, connected subsets of W is a biconnected set B, then the problem concerning the existence of a dispersion point is solved in the negative. For if $\bar{B}$ is an indecomposable continuum, B cannot contain a cut-point. See P. M. Swingle, "Two Types of Connected Point Sets," *Bulletin of the American Mathematical Society*, Vol. 37 (1931).

‡ E. Zermelo, "Beweis dass jede Menge wohlgeordnet werden kann," *Mathematische Annalen*, Vol. 59 (1904), pp. 514-516.

§ For definitions and properties see Z. Janiszewski and C. Kuratowski, "Sur les continus indécomposables," *Fundamenta Mathematicae*, Vol. 1, pp. 217-222.

a point of $T + a + b$. As there exist c composants * of W, it follows that $C_i \times W$ is of power c. Also it is known that P_i is of power c.

Since $c \times c = c$, it follows that $m \times c = c$. Then by lemma A there exist mutually exclusive sets $K_1, K_2, \cdots, K_h$, where the power of h is m, and mutually exclusive sets Q_1 and Q_2 of W, where (1) $K_f \times Q_g = 0$ $(f = 1, 2, \cdots, h;\ g = 1, 2)$, (2) each K_f contains a point of every C_i of (C_i), and each Q_g contains a point of every P_i of (P_i).

Let all the points of W, which are not contained either in a K_f or in Q_2, be added to K_1, obtaining a new set K. Then each of the sets K, $K_2 + Q_2$, K_3, $K_4, \cdots, K_h$ is punctiform, for Q_1 contains a point, which is contained in none of these sets except K, of each perfect set of W; and Q_2 contains a point, which is not contained in K, of each perfect set of W. As each of these sets contains a point of every C_i of (C_i), they are connected.† And it is evident that each of these sets is dense in W. Therefore it is seen that the theorem is true.

Examples of indecomposable continua, every proper subcontinuum of which is a simple continuous arc, are well known. Hence there exist indecomposable continua no proper subcontinuum of which contains a punctiform connected subset. For such continua the following corollary is seen to be true.

COROLLARY A. *Let W be a bounded indecomposable continuum, lying in a euclidean space, no proper subcontinuum of which contains a punctiform connected subset. Let m be any cardinal number greater than one and less than c, the power of the linear continuum. Then W is the sum of m mutually exclusive, widely connected subsets, each of which is dense in W.*‡

By a proof similar to the above it is seen that *a bounded domain D, lying in a euclidean space, is the sum of m mutually exclusive, punctiform, connected subsets, each dense in D, where m is any cardinal number greater than one and less than c, the power of the linear continuum.*§

A domain D, lying in a euclidean space of n dimensions, may be the sum of m biconnected subsets, which are mutually exclusive except for one point and are each dense in D, where m is any cardinal number greater than n and less than c, the power of the linear continuum. For consider the following example.

---

* S. Mazurkiewicz, "Sur les continus indécomposables," *Fundamenta Mathematicae*, Vol. 10, pp. 305-310.

† B. Knaster and C. Kuratowski, *Fundamenta Mathematicae*, Vol. 2, pp. 233-234, Theorem 37.

‡ A connected set will be said to be *widely connected* if every proper connected subset is dense in it. See P. M. Swingle, "Two Types of Connected Sets," *loc. cit.*

§ See F. Bernstein, *Leipziger Berichte*, Vol. 60 (1908).

Example A. Consider in the euclidean plane the straight line interval bc and the point a not on the line bc. It will be shown that the domain bounded by the triangle abc is the sum of three biconnected sets, having only a in common, which is a dispersion point of each. The interval bc is the sum of three mutually exclusive subsets U_1, U_2, and U_3, each of power c and each dense in bc. Let W_i $(i = 1, 2, 3)$ be the set of straight line intervals each of which join a to a point of U_i. Then by a proof, similar to that of theorem A, it is seen that W_i is the sum of two punctiform connected subsets, B_{1i} and B_{2i}, having only a in common and each dense in W_i. But as a is a dispersion point of B_{ji} $(j = 1, 2)$, B_{ji} is biconnected. Hence the interior of the triangle plus its boundary is the sum of the three biconnected subsets, $B_{11} + B_{12}$ $B_{21} + B_{23}$, and $B_{22} + B_{13}$, having only a in common, a being the dispersion point of each of these sets, and each of these sets is dense in the interior of the triangle. It is seen readily that the above method can be modified to show that the interior of the triangle abc is the sum of three biconnected subsets, having only one point in common, where each of these biconnected subsets is contained in the interior of the triangle. And the above method easily is modified to show that either of these sets is the sum of m such biconnected subsets, even if $n > 2$.

THEOREM 5. *An n-dimensional compact metric space, S, is never contained in the sum of n biconnected sets.*

Assume that S is contained in the sum of the n biconnected sets B_i $(i = 1, 2, \cdots, n)$. As a biconnected set does not contain two distinct connected subsets, each set of $n + 1$ mutually exclusive n-dimensional spheres of S contains one such sphere, S_n say, which does not contain a connected subset of any B_i. But S_n contains an n-dimensional Cantorian multiplicity C_n.* Say for example that B_n contains two distinct points of C_n, as one of the B_i's will. Then there exists a subcontinuum of S_n, and so a closed subset K_{n-1} of $C_n - B_n$, which separates B_n.† And as C_n is a Cantorian multiplicity,‡ K_{n-1} must be of dimension $n - 1$. But as a closed subset of a compact space is itself a compact space, K_{n-1} contains an $(n-1)$-dimensional Cantorian multiplicity C_{n-1} as a subset, having the property that it is contained in $n - 1$ of the biconnected sets B_i. Hence, proceeding as above, by induction it follows that there exists a one-dimensional Cantorian multi-

---

* K. Menger, *Dimensiontheorie* (1928), p. 217, Theorem 8.

† C. Kuratowski and B. Knaster, "Sur les ensembles connexes," *Fundamenta Mathematicae*, Vol. 2, pp. 233-235, Theorem 37.

‡ P. Urysohn, "Mémoire sur les multiplicités Cantoriennes," *Fundamenta Mathematicae*, Vol. 7, p. 124, Definition.

plicity C_1, which is contained in one of the biconnected sets B_i. Hence, as C_1 is a continuum contained in a punctiform set, a contradiction has been obtained.

COROLLARY 3. *If D is a domain of a compact metric space containing a continuum C, which is the sum of two biconnected sets, then $\overline{D - C} = \overline{D}$.*

This is true because no point of C is an interior point of C.

It has been shown that a domain in the plane is not the sum of two biconnected sets. It will be shown now that there is a continuum in the plane, which is the sum of two mutually exclusive biconnected sets. Let, in the euclidean plane, T be a Cantor ternary set on the straight line interval bc, and let a and d be two points not on the line bc, where a and d lie on opposite sides of this line. Let W be the continuum obtained by joining the points of T to each of the points a and d by straight line intervals. Then $W - d$ is seen to be, by a proof similar to that of theorem A, the sum of two biconnected sets B_1 and B_2, having a in common. Therefore W is the sum of the two mutually exclusive biconnected subsets B_1 and $B_2 - a + d$. It is to be noted that W is a compact metric space of dimension one. Also each of the biconnected subsets is of dimension one.

THEOREM 6. *An n-dimensional euclidean space, E_n, does not contain an n-dimensional biconnected subset.*

Assume that E_n contains an n-dimensional biconnected subset B. Then by definition of an n-dimensional space, there exists an n-dimensional sphere S_n whose boundary S_{n-1} contains an $n-1$ dimensional punctiform subset of B. And similarly S_{n-k} $(k = 1, 2, \cdots, n-2)$ will have as boundary S_{n-k-1}, containing an $n-k-1$ dimensional punctiform subset of B. Hence by induction there exists a one-dimensional sphere S_1, which contains a punctiform one-dimensional subset of B. Thus a contradiction is obtained.*

It readily is seen that, if n is greater than one, an n-dimensional euclidean space may contain a k-dimensional biconnected subset, where k is any integer from one to $n - 1$.

3. *Euclidean spaces as the sum of biconnected sets.* In this section it will be shown that a euclidean space can be the sum of mutually exclusive biconnected subsets.†

---

* P. Urysohn, *loc. cit.*, p. 76.

† See P. M. Swingle, "Generalizations of Biconnected Sets," *American Journal of Mathematics*, Vol. 53 (1931), Problem 2, p. 392. Parts of this problem are not answered in this section; for example, whether a euclidean plane can be the sum of c mutually exclusive biconnected subsets, where c is the power of the linear continuum, remains unanswered.

In Example A of the previous section it was shown, by means of Zermelo's postulate, that, if T was the interior of a triangle bae in E_2, then $\bar{T}$ is the sum of m biconnected subsets, where m is any cardinal number greater than 2 but less than c, mutually exclusive except for one point which is a dispersion point. Each of these biconnected subsets are dense in T.

It will be shown now that *the euclidean plane E_2 is the sum of m mutually exclusive biconnected subsets, where m is any integer greater than* 2. For let S be the points of E_2, and let e be the center of a regular polygon P of m sides. Let on the circle, with center e and radius one, U_j $(j=1,2,3)$ be three mutually exclusive punctiform point sets of power c and each dense on this circle, whose sum is this circle. Let W_j be the points on the straight line intervals, from e to infinity, passing from e through the points of U_j. Then by a proof, similar to that of theorem A, it is seen that S is the sum of m biconnected subsets, $G_1, G_2, \cdots, G_m$, mutually exclusive except for the common dispersion point e and each dense in S.

Let a_i $(i=1,2,\cdots,m)$ be the vertices of P. Each triangle $ea_{i+1}a_i$, where $m+1=1$, is the sum of m biconnected subsets $B_{(i+1)g}+a_{i+1}$ $(g=1,2,\cdots,m)$, where the $B_{(i+1)g}$'s are mutually exclusive and the point a_{i+1} is the dispersion point of each biconnected subset. Furthermore $(B_{(i+1)(i+1)}+a_{i+1})\times ea_i$ can be taken as vacuous, while each of the other biconnected sets is dense on ea_i. Let $G_{(i+1)e}$ be the points of G_{i+1} exterior to P and let

$$N_{i+1}=B_{(i+1)1}+B_{(i+1)2}+\cdots+B_{(i+1)m}+a_{i+1}+G_{(i+1)e}.$$

Then N_{i+1} is a biconnected set with dispersion point a_{i+1} and dense in S. Hence the euclidean space E_2 is the sum of the m mutually exclusive biconnected subsets N_{i+1}.

That E_n, $n>2$, *is the sum of m mutually exclusive biconnected subsets, where m is any integer greater than n,* is shown in a similar manner.

A similar proof will show that E_n *is the sum of a countable infinity of mutually exclusive biconnected subsets.* For consider E_2. Take the regular polygon with five sides. Let $a_6, a_7, \cdots$ be an infinite sequence of distinct points on a_5a_1, with sequential limit point a_1. The proof then follows as above.

That E_3 *is the sum of c mutually biconnected subsets, where c is the power of the linear continuum,* follows from the fact that E_3 is the sum of c mutually exclusive planes, each of which is the sum of a finite number of biconnected subsets.

That E_2 is the sum of a countable infinity of mutually exclusive biconnected subsets, each with a dispersion point but not each dense in E_2, can be shown by means of the following interesting example of a biconnected set.

In E_2 take a cartesian coördinate system. By a biconnected set

$$((i,j)(i,j+1)(i+1,j)),$$

or similar set, will be meant one dense, and contained, in a triangle having these points as vertices and with the point $(i, j+1)$ as dispersion point of this biconnected set. It will be further understood that this biconnected set does not contain any point, excepting its dispersion point, on the line interval from (i,j) to $(i,j+1)$ nor on the line interval from $(i+1,j)$ to $(i,j+1)$, but that it is dense on the interval from (i,j) to $(i+1,j)$, excepting that the biconnected set $((0,0)(0,1)(1,0))$ also has nothing common with the interval from $(0,0)$ to $(1,0)$. Let i and j take on all values $\pm 1, \pm 2, \pm 3, \cdots, 0$.

Let $T_{(+j)}$ $(j=1,2,\cdots)$ be the set of all the points, excepting the dispersion point, of the following infinite set of biconnected sets:

$$\begin{array}{ll}
((i,j)(i+1,j)(i+1,j-1)) & (i=0,1,2,\cdots), \\
((i,j)(i-1,j)(i,j-1)) & (i=0,-1,-2,\cdots), \\
((i,j-1)(i+1,j-1)(i,j)) & (i=1,2,\cdots), \\
((i,j-1)(i-1,j-1)(i-1,j) & (i=0,-1,-2,\cdots)
\end{array}$$

and

$$((0,j-1)(0,j)(1,j-1)).$$

To obtain $T_{(-j)}$ $(j=1,2,\cdots)$ rotate $T_{(+j)}$ 180° about the X-axis and translate the resulting set, parallel to itself, through an X distance of -1; for example $((0,0)(0,1)(1,0))$ will become $((-1,0)(-1,-1)(0,0))$. However let this last set be dense on the interval from $(-1,0)$ to $(0,0)$.

Let g be the point $(0,1)$. Then

$$g+T_1+T_2+\cdots+T_{-1}+T_{-2}+\cdots=B$$

is a biconnected set everywhere dense in E_2, connected im kleinen at g and having g as dispersion point. It has the interesting property that only one of the quasi-components of $B-g$ contains a point outside of the triangle $((0,0)(0,1)(1,0))$.

It is easy to see that if S is the set of all the points of E_2, excepting the countable infinity of points of E_2 which are dispersion points of the above biconnected sets, and if $g_1=(0,2)$ and $g_2=(0,3)$, then $S+g+g_1+g_2$ is the sum of three mutually exclusive biconnected sets, each dense in E_2 and having dispersion points g, g_1, and g_2 respectively. And it is further seen that E_2 is the sum of a countable infinity of mutually exclusive biconnected sets, three of which are dense in E_2 and the remaining ones are each con-

tained in one of the triangles used above. Similar results can be obtained for E_n, $n > 2$.

From the results of the last two sections the following theorem now will be proved.

THEOREM B. *In order that a euclidean space E be n dimensional, n a positive integer greater than one, it is necessary and sufficient that E be the sum of $n+1$ but not of n biconnected subsets.**

The condition is necessary. For by theorem 5, E_n is not the sum of n biconnected subsets. And in this section it is shown, by means of Zermelo's postulate, that E_n is the sum of $n+1$ mutually exclusive biconnected subsets.

The condition is sufficient. For E is the sum of $n+1$ biconnected subsets by hypothesis, and so is not k dimensional, k greater than n, for E_k contains an E_{n+1}, which by theorem 5 is not contained in the sum of $n+1$ biconnected sets. And it is not a k dimensional euclidean space, k less than n, for such a space could be the sum of n biconnected subsets, excepting for $k=1$ which would not contain a biconnected set. Hence E must be n dimensional.

OHIO STATE UNIVERSITY.

* For an interesting similar result see P. Urysohn, *Fundamenta Mathematicae*, Vol. 8, Corollary, p. 345; see also Karl Menger's *Dimensiontheorie*, p. 113 and p. 120

A NOTE ON SPACES HAVING THE S PROPERTY.

By Gordon T. Whyburn.

1. A metric space D is said to have property S * provided that for each $\epsilon > 0$, D is the sum of a finite number of connected sets each of diameter $< \epsilon$. Since all such spaces have only a finite number of components, little generality is lost by supposing them connected. If N is any subset of a metric space D and ϵ is any positive number, we shall denote by $T_\epsilon(N)$ the set of all points x of D which can be joined to some point p of N by a chain of connected subsets $L_1, L_2, \cdots, L_n$ of D such that for each i, $\delta(L_i) < \epsilon/2^i$, $L_1 \supset p$, $L_n \supset x$, and any two successive sets (links) L_i and L_{i+1} have at least one common point. Such a chain will be called a chain of type T_ϵ or simply type T.

2. Theorem. *If the metric space D has property S, then every set $T_\epsilon(N)$ has property S.*

For let δ be any positive number, and let us choose an integer k such that $\sum_k^\infty \epsilon/2^i < \delta/4$. Let E be the set of all points in $T_\epsilon(N)$ which can be joined to N (i.e., to some point of N) by a chain of type T which has at most k links. Let us express D as the sum of a finite number of connected sets each of diameter $< \epsilon/2^{k+1}$; and of these sets, let $Q_1, Q_2, \cdots, Q_n$ be the ones which contain at least one point of E. Then we have $E \subset \sum_1^n Q_i$. Now for each i, $Q_i \subset T_\epsilon(N)$; for Q_i contains a point x of E, and x can be joined to N by a chain $L_1, L_2, \cdots, L_r$ of type T having k links or less; and since $\delta(Q_i) < \epsilon/2^{k+1}$, therefore $[L_1, L_2, \cdots, L_r, Q_i]$ is a chain of type T; and hence $Q_i \subset T_\epsilon(N)$. For each i, $1 \leqq i \leqq n$, let W_i be the set of all points of $T_\epsilon(N)$ which can be joined to some point of Q_i by a connected subset of $T_\epsilon(N)$ of diameter $< \delta/4$. Then for each i, W_i is a connected subset of $T_\epsilon(N)$ of diameter $< \delta$. It remains only to show that $T_\epsilon(N) \subset \sum_1^n W_i$. To this end let x be any point of $T_\epsilon(N)$ and let $L_1, L_2, \cdots, L_m$ be a chain of type T joining x to N. Obviously we need only consider the case in which x does not belong to E, and in this case $m > k$. Then since $L_k \subset E$, it

* See Sierpinski, *Fundamenta Mathematicae*, Vol. 1, p. 44; and R. L. Moore, *ibid.*, Vol. 3, p. 232.

follows that for some j, $L_k \cdot Q_j \neq 0$; and since $\sum_k^\infty \epsilon/2^i < \delta/4$, it follows that $\delta(\sum_k^m L_i) \leq \sum_k^m \delta(L_i) < \delta/4$. Hence $\sum_k^m L_i$ is a connected subset of $T_\epsilon(N)$ of diameter $< \delta/4$ which joins x to a point of Q_j. Therefore $x \subset W_j$, and our theorem is proved.

3. Corollary 1. *Any metric space D having property S is the sum of a finite number of arbitrarily small connected subsets each having property S. Furthermore these subsets may be chosen either as open sets or as closed sets.*

For let δ be any positive number, let $\epsilon = \delta/3$, and let $D = \sum_1^n D_i$, where each D_i is connected and of diameter $< \epsilon$. Then, for each i, $T_\epsilon(D_i)$ is connected and of diameter $< \delta$ and clearly $D = \sum_1^n T_\epsilon(D_i)$. Now the sets $[T_\epsilon(D_i)]$ themselves are open; and since it is true that if a set E has property S, so does every set E_0 such that $E \subset E_0 \subset \bar{E}$, it follows that the sets $[\overline{T_\epsilon(D_i)}]$ have property S, and of course they are closed.

From this corollary it follows that in any metric space having property S there exists a monotone decreasing *fine* subdivision into connected sets. In other words, we can subdivide such a space into a finite number of connected sets each having property S and being of diameter < 1; then we can subdivide each of these sets into a finite number of connected sets of diameter $< 1/2$, and so on indefinitely.

4. Corollary 2. *Any point p of a metric space D having property S is contained in an arbitrarily small connected open set (region) which has property S.*

To see this we have only to take $N = p$, and then the set $T_\epsilon(p)$ is the desired region. Since $\overline{T_\epsilon(p)}$ also has property S and hence is locally connected (because any set having property S is locally connected), we have shown that p is contained in an arbitrarily small region whose closure is locally connected.

5. For any point p of a metric space D and for each $\epsilon > 0$ let $R_\epsilon(p)$ be the set of all points x of D which can be joined to p by a chain of connected sets $L_1, L_2, \cdots, L_n$ in D such that $L_1 \supset p$, $L_n \supset x$, any two successive links have a common point, and for each i, $\delta(L_i) < \epsilon - \sum_1^i \delta(L_j)$.* Then by

* See my paper "Concerning S-Regions in Locally Connected Continua," appearing in *Fundamenta Mathematicae.*

essentially the same argument as given in § 2 it is shown that if D has property S, every set $R_\epsilon(p)$ likewise has property S.

Now let us suppose D has property S. For any ordered pair of points x, y in D let us denote by $\rho^+(x, y)$ the greatest lower bound for the aggregate of numbers $[\epsilon]$ such that $y \subset R_\epsilon(x)$, and let D^+ denote the space whose points are the same as those of D but in which distances are defined by the function ρ^+. It is readily shown that ρ^+ satisfies all the axioms for a metric except the symmetry axiom $[\rho(x, y) = \rho(y, x)]$, which may not be satisfied; and thus D^+ is a quasi metric space. Furthermore, D^+ is homeomorphic with D; and the spherical neighborhoods in D^+ are exactly the sets $R_\epsilon(p)$ (thus they are connected and have property S). This naturally raises the question of the existence of *some* distance function ρ which satisfies *all* the axioms for a metric and which gives these properties to the new space. In other words, *Is every metric space D which has property S homeomorphic with some metric space D' in which all spherical neighborhoods are connected and have property S?*

A positive solution to this problem even in the case of compact connected spaces D having property S (i. e., compact locally connected continua) would yield a very convenient and useful machinery with which to study such spaces.

THE JOHNS HOPKINS UNIVERSITY.

REMARKS ON THE RECURRENT BEHAVIOR OF THE CHARACTERISTICS ON A TORUS.

By Aurel Wintner.

A paper of the author on the Levi-Civita problem of mean motion (to appear in the *Annali di Matematica*) seems to lead to some contradictions with an important discovery by Denjoy, recently published in the *Comptes Rendus*. On combining the theorem of Denjoy with other results it will be shown in the present note that, in reality, those apparent contradictions do not exist at all.

Let $f(\vartheta, t)$ be a real-valued continuous function defined in the real (ϑ, t)-plane and possessing the period 1 with respect to both ϑ and t. Let T denote the closed orientable surface of genus one representing the fundamental domain $0 \leqq \vartheta \leqq 1$, $0 \leqq t \leqq 1$. The function f should possess the additional property that the differential equation

$$(1) \qquad d\vartheta/dt = f(\vartheta, t)$$

admits, for a given real initial condition $\vartheta(t_0) = \vartheta_0$, only one solution $\vartheta = \vartheta(t)$, depending on the initial condition in a continuous manner. This additional condition which will be designated as the condition Γ is, for instance, fulfilled if $\partial f/\partial\vartheta$ exists and is continuous. If every solution

$$(2) \qquad \vartheta = \vartheta(t)\,; \; -\infty < t < +\infty$$

of (1) represents an everywhere-dense point-set on the torus T we shall say that the function f is of the ergodic type.

It is known since Poincaré * that if f fulfills the condition Γ then there exists for every solution (2) of (1) a constant μ and a *bounded* function $\omega(t)$ so that

$$(3) \qquad \vartheta(t) = \mu \cdot t + \omega(t)$$

where the "mean motion" or "rotation number" μ is independent of the special choice of the integration constant $\vartheta(t_0) = \vartheta_0$ and is, therefore,

* Cf. H. Kneser, "Reguläre Kurvenscharen auf den Ringflächen," *Mathematische Annalen*, Vol. 91 (1924), pp. 135-154, where references are also given. The proof of a lemma occurring in the demonstration has been simplified by J. Nielsen, "On topologiske Afbildninger af en Jordankurve paa sig selv," *Matematisk Tidsskrift*, B, No. 3 (1928), p. 39-46. A very simple direct proof of (3) has recently been given by A. Weil, "On Systems of Curves on a Ring-Shaped Surface," *The Journal of the Indian Mathematical Society*, Vol. XIX (1931), pp. 109-114.

uniquely determined by the double-periodic function f. It may be mentioned that the number μ which is obtained by an infinite process of iteration cannot be less than the minimum or larger than the maximum of the double-periodic function $f(\vartheta, t)$. In fact, on substituting (3) in (1) and using the fact that $\omega(t)$ is a bounded function, we obtain that the time-average

$$\lim_{t=+\infty} \int_0^t f(\vartheta(\tau), \tau)\, d\tau/t$$

exists and is equal to μ.

On account of the boundedness of the remainder term $\omega(t)$ one may expect that f is of the ergodic type provided that μ is irrational. This is, however, not necessarily the case, even if $\partial f/\partial\vartheta$ exists and is continuous. On refining the methods of his earlier paper on this problem Kneser has shown in a note to appear in the *Annali di Matematica* that if the condition Γ is fulfilled and if μ is irrational the function f is then and only then of the ergodic type if the remainder term $\omega(t)$ is, for all values of the integration constant $\vartheta(t_0) = \vartheta_0$, an almost-periodic function in the sense of Bohr. On the other hand, it has been conjectured by Poincaré (1885) that f is necessarily of the ergodic type if it is regular analytic in the real domain (provided that μ is irrational). The correctness of this conjecture which is of great importance for the dynamics of Birkhoff * has been proved by Denjoy † in his fundamental note mentioned above. Denjoy shows that even the condition of analyticity is superfluous, viz., it is sufficient that a certain "real" condition is fulfilled which is only a little stronger than the non-sufficient condition Γ.

We combine now these results with the following fact, explicitly pointed out also by Bohl: ‡ if the condition Γ is fulfilled and μ is irrational the function f is then and only then of the ergodic type if there exists, for any value of the integration constant $\vartheta(t_0) = \vartheta_0$, a continuous double-periodic function

$$\rho = \rho(\xi, \eta) = \rho(\xi + 1, \eta) = \rho(\xi, \eta + 1),$$
$$-\infty < \xi < +\infty, \quad -\infty < \eta < +\infty$$

for which

---

* Cf., for instance, G. D. Birkhoff, "Surface Transformations and their Dynamical Applications," *Acta Mathematica*, Vol. 43 (1920), p. 79 etc., and "On the Periodic Motions of Dynamical Systems," *Acta Mathematica*, Vol. 50 (1927), pp. 359-379.

† A. Denjoy, "Sur les caractéristiques à la surface du tore," *Comptes Rendus des Séances de l'Académie des Sciences*, Vol. 194 (1932), pp. 830-833.

‡ P. Bohl, "Ueber die hinsichtlich der unabhängigen und abhängigen Variabeln periodische Differentialgleichung erster Ordnung," *Acta Mathematica*, Vol. 40 (1916), p. 321 etc.

$$(4) \qquad \omega(t) = \rho(\mu \cdot t, t).$$

Since any function $\omega(t)$ possessing a representation (4) is almost-periodic in the sense of Bohr * it follows from the result of Kneser mentioned above that if the remainder term $\omega(t)$ is at all almost-periodic then it possesses only two linearly independent frequencies (viz., μ and 1). On the other hand, on combining the remark of Bohl with the theorem of Denjoy it follows that $\omega(t)$ is in the analytic case necessarily almost-periodic, provided μ is irrational.

Let $a(t)$, $b(t)$, $c(t)$, $d(t)$ be four functions which are real-valued and regular analytic for $-\infty < t < +\infty$ and possess the period 1. On denoting by

$$(5) \qquad x = x(t), \ y = y(t); \quad x(t)^2 + y(t)^2 > 0$$

a real solution of the differential equations

$$(6) \qquad dx/dt = a(t)x + b(t)y, \quad dy/dt = c(t)x + d(t)y$$

and placing

$$(7) \qquad x(t) = r(t) \cos(\vartheta/2\pi), \quad y(t) = r(t) \sin(\vartheta/2\pi) \text{ where } r(t) > 0$$

one obtains from (6), according to Levi-Civita,† the single differential equation

$$(8) \quad d\vartheta/dt = 2\pi\{c(t) \cos^2(\vartheta/2\pi) + [d(t) - a(t)] \cos(\vartheta/2\pi) \sin(\vartheta/2\pi) - b(t) \sin^2(\vartheta/2\pi)\}$$

which is, of course, a special case of (1) and satisfies the regularity condition of Denjoy. In the note mentioned above I pointed out that $\omega(t)$ is, in the special case (8) of (1), not necessarily almost-periodic. It follows, therefore, from the theorem emphasized at the end of the previous paragraph that μ cannot be, in these cases, an irrational number. On the other hand, in the presentations of the problem ‡ it is usually shown that $\omega(t)$ is, even under the single condition Γ, not only almost-periodic but also periodic if μ is a rational number. Accordingly $\omega(t)$ is then in the case (8) *always* almost-periodic. This is, however, in contradiction with the following examples:

Suppose that the four coefficients of (6) are not only periodic but even

---

* H. Bohr, "Zur Theorie der fastperiodischen Funktionen. II.," *Acta Mathematica*, Vol. 46 (1925), p. 134 etc.

† T. Levi-Civita, "Sur les équations à coefficients périodiques et sur le mouvement moyen du noeud lunaire," *Annales de l'École Normale Supérieur* (3), Vol. 28, pp. 325-376.

‡ Cf., for instance, A. Weil, *loc. cit.*

independent of t. Equation (8) is also in this case a special case of equation (1). Suppose further that the characteristic numbers of the matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

are real and that the determinant of the matrix is positive. There are three cases possible, according as the characteristic numbers and the elementary divisors are simple or double. However, in all three possible cases the origin of the (x, y)-plane is a node (noeud, Knoten),* i. e., the curve (5) which is in general not a straight line possesses in the point $x = y = 0$ a tangent. We conclude from (7) that the function $\vartheta(t)$ possesses for † $\lim t = \infty$ a finite limit α but is not equal to this constant α for all values of t (provided the curve is not a straight line which may be excluded). Since the limit α exists it follows from (3) that $\vartheta(t)$ coincides with the bounded function $\omega(t)$. Furthermore, since $\vartheta(t)$ is not constant and possesses for $\lim t = \infty$ a limit it follows that $\omega(t)$ is *not* an almost-periodic function.

This paradox is cleared up as follows: On reading over the usual demonstrations for the fact that $\omega(t)$ is a periodic function for rational values of μ one may observe that the proof breaks down if the rational rotation number μ is equal to zero as is the case in our examples.

THE JOHNS HOPKINS UNIVERSITY.

---

* Cf., for instance, E. Kamke, "Differentialgleichungen reeller Funktionen," Leipzig, 1930, p. 200 etc. Intuitive pictures for the different kinds of nodes are given by O. Perron, "Ueber die Gestalt der Integralkurven einer Differentialgleichung erster Ordnung in der Umgebung eines singulären Punktes. (Erster Teil)," *Mathematische Zeitschrift*, Vol. 15 (1922), p. 123.

† It depends on the matrix of the coefficients whether $\infty = +\infty$ or $\infty = -\infty$.

SOME REMARKS ON THE JOIN OF TWO COMPLEXES AND ON INVARIANT SUBSETS OF A COMPLEX.

By EGBERT R. VAN KAMPEN.

In this paper we consider some miscellaneous properties of combinatorial topology mostly found during the preparation of a course on this subject. The first three sections deal with the homology characters of a join; it is found useful to introduce homology characters in a combinatorial way for a kind of open complex slightly more general than is customary. After that we prove some simple properties of invariant subsets of complexes, soon concentrating on the problem of dividing any complex into the smallest invariant subsets that can be found with the methods now at the disposal of combinatorial topology. To do this we have to prove a number of properties on the invariance in a complex of the invariants of the neighborhood complexes of the different simplexes of the complex, using some of the properties of joins proved previously.

1. As the cells of the join $(K_1 \cdot K_2)$ and the product $(K_1 \times K_2)$ of two complexes (K_1) and (K_2) as well as their incidence relations are completely determined by K_1 and K_2 it is a matter of calculation to find their homology characters when K_1 and K_2 are given. Superficially the computation seems easier for the product because in the product there is one $(p+q)$-cell corresponding to every pair formed by a p-cell of K_1 and a q-cell of K_2 and no other cell; the boundary of the $(p+q)$-cell corresponds to all the pairs which can be formed by replacing the p- or the q-cell by its boundary cells. The structure of a join is slightly more complicated. This induced A. B. Brown * to reduce the computation for $K_1 \cdot K_2$ to that for $K_1 \times K_2$ by means of a relatively complicated reasoning on different kinds of chains on the join; he does not find any result on torsion numbers. However, when we introduce formally a (-1)-dimensional simplex into all the relevant complexes forming the boundary of all vertices and such that its join with any cell is that cell, then there is in the join a $(p+q+1)$-simplex corresponding to every pair formed by a p-simplex of K_1 and a q-simplex of K_2; with analogous boundary relations as for the product. The result for the product † can now be transcribed for the join without any additional calculation.

* "On the Join of two Complexes," *Bulletin of the American Mathematical Society*, Vol. 37 (1931), p. 417.

† See for instance E. R. van Kampen, "Die kombinatorische Topologie und die Dualitätssätze," Van Stockum, *Den Haag*, p. 37.

THEOREM I. *Suppose the homology characters of K_1, K_2 and $K_1 \cdot K_2 = K$ to be computed after the introduction of a (-1)-simplex which is on the boundary of every 1-simplex (that means putting R^0 equal to the number of components minus one), then we have for the homology characters of K:*

$$R^p(K) = \sum_{k+l+1=p} R^k(K_1) \cdot R^l(K_2).$$

The numbers $\tau_i^p(K)$ are the elementary divisors of a matrix of which the elements outside of the diagonal are zero while the elements in the diagonal are:

$R^k(K_2)$ *times the numbers* $\tau_i^{p-k-1}(K_1)$, $\quad (k = 0, \cdots, p-1)$;
$R^k(K_1)$ *times the numbers* $\tau_i^{p-k-1}(K_2)$, $\quad (k = 0, \cdots, p-1)$;

the greatest common divisors of all pairs

$\tau_i^k(K_1)$, $\tau_j^{p-k-1}(K_2)$, $\quad (k = 1, \cdots, p-2)$,
and of all pairs $\tau_i^k(K_1)$, $\tau_j^{p-k-2}(K_2)$, $\quad (k = 1, \cdots, p-3)$.

2. We are now going to define homology characters of L modulo M, where L and M are sums of the interiors of certain simplexes of a complex K, L containing M, and M containing its limit-points insofar as these belong to L.* As the case considered here is barely more general than that worked out in detail by S. Lefschetz in his Topology, we give only the definitions, none of the proofs.

The homology characters of L modulo M are defined as the homology characters of $\bar{L}$ modulo $\bar{M}$ where $\bar{L}$ is a subcomplex of a subdivision of K approximating L and $\bar{M}$ is the intersection of $\bar{L}$ and M, which is also a complex. For $\bar{L}$ we take the point set obtained from K by applying the following process to every simplex of K of which the interior is not on L: Construct a copy U of the neighborhood complex of that simplex in a very small neighborhood of one of its interior points and take away from K the join of U and the simplex, except the join of U and the boundary of the simplex. U must be taken so small, that the parts taken away for two simplexes do not have a common point when the simplexes are not incident. $\bar{L}$ is a subcomplex of a subdivision of K; any chain on L of which some part is on M can be deformed into a chain of $\bar{L}$, the part on M remaining on M and finally moving into $\bar{M}$. The homology characters of L modulo M are invariant under transformation of K into another complex carrying L and M into sums of interiors of simplexes. This follows from the following:

* If this is not true from the beginning the logical approximation for M is exactly the inclusion of those limit-points.

THEOREM 2. *The homology characters of L modulo M can be defined by means of singular chains on L neglecting those contained in M.*

3. We want to find point-sets L_3 and M_3 on $K_1 \cdot K_2 = K_3$ such that the formulas of theorem 1 give the homology characters of L_3 modulo M_3 as functions of those of L_1 modulo M_1 (on K_1) and L_2 modulo M_2 (on K_2). This problem can be reduced to the following: What is the point-set L_3 approximated on K_3 by the join $\bar{L}_3$ of the point-sets $\bar{L}_1$ and $\bar{L}_2$ approximating L_1 and L_2 and what is the point-set M_3 approximated by $\bar{M}_3$ which is the sum of those simplexes of $\bar{L}_3$ corresponding to a pair of simplexes of which at least one is in M_1 or M_2. For the computation as for the product and so at the same time theorem 1 can be applied on $\bar{L}_3$ modulo $\bar{M}_3 \cdot L_3$ turns out to be the set of all segments between points of L_1 and of L_2 on K_3; we can write $L_3 = L_1 \cdot L_2$. The approximation of L_3 by $\bar{L}_3$ is of a slightly different kind from that described in section 2. With the same notation $M_3 = L_1 \cdot M_2 + M_1 \cdot L_2$. The introduction of the (-1)-simplex in K_i does not have any influence on the homology characters of L_i modulo M_i provided M_i is not empty because in that case the (-1)-simplex has to be included in M_i. The repetition of Theorem 1 in its generalized form seems superfluous.

4. We base the treatment of invariant subsets on the following definition: *An invariant subset of a complex K is a subset which is transformed into itself by any transformation of K into itself or the sum of a certain number of components of such point-sets.* Combinatorial invariants of invariant subsets of K are combinatorial invariants of K. A minimal invariant subset of K is an invariant subset not containing an invariant subset different from itself. The interesting problem of finding all minimal invariant subsets of a complex is beyond the scope of the present methods of combinatorial topology.

The minimal invariant subset containing a point P of K is the component of P in the set of points of K into which P can be transformed by topological transformations of K into itself.

As any interior point of a simplex of K can be transformed into all other interior points of the same simplex, any invariant subset is the sum of the interiors of a certain number of simplexes of K. The closure of an invariant subset of K is an invariant subcomplex of K.

THEOREM 3. *There is an isotopic deformation of K into itself under which a point of any minimal invariant subset of K goes into another point of the same subset.*

To prove this we note first that there is such a deformation moving any point interior to a simplex into any other point interior to the same simplex;

second that any point of the invariant subset can be transformed into any other point of the subset. From this it follows that any point of the subset can be deformed into any point of a small neighborhood of the invariant subset and that this neighborhood can be extended to contain all interiors of simplexes belonging to the invariant subset on which, or on the boundary of which P lies. Finally it is obviously possible to pass from any point to any other point of the subset in a finite number of steps all completely contained in one of the enlarged neighborhods.

THEOREM 4. *If a simplex S^p of K^n is contained in the closure of an invariant subset M of K^n, then the point-set N, corresponding in the complement U^{n-p-1} of S^p in K^n to M, is an invariant subset of U^{n-p-1}.*

This follows from the fact that the interior points of the segments between interior points of S^p and the points of a minimal invariant subset of U^{n-p-1} all belong to the same invariant subset of K^n. It is sufficient to construct an isotopic deformation of K^n into itself moving part of such a segment into part of an arbitrary other segment, supposing that the two segments pass through the same point P of S^p and for instance through the points Q and R of U^{n-p-1}. According to theorem 3 there is an isotopic deformation of U^{n-p-1} into itself moving R into Q; this determines an isotopic deformation of the join U^{n-1} of U^{n-p-1} and the boundary of S^p into itself still moving R into Q. U^{n-1} is the neighborhood complex of P in K^n. A certain neighborhood of P, the join of P and U^{n-1}, can be formed by means of a set of copies of U^{n-1} depending on a parameter moving in an interval, however to one end-point of the interval there only corresponds the point P. We construct the isotopic deformation of K^n in such a way that under any transformation of that deformation all those copies of U^{n-1} are transformed into themselves by means of one of the transformations of the deformation of U^{n-1}. The copies near P are transformed by means of the final transformation of that deformation, the copies near the boundary of our neighborhood of P and at the same time all points outside of that neighborhood are transformed by means of the identical transformation, the copies in between those two groups are transformed by means of all the transformations of our deformation of U^{n-1} distributed continuously. The deformation of K^n is executed by moving this intervening portion continuously till it arrives at P. It is easy to see that this deformation satisfies the condition we wanted so that we have proved theorem 4.

All theorems which we are going to give with the object of finding smaller invariant subsets of a complex except theorem 10 will satisfy a certain converse of theorem 4: to any invariant point-set of a complement found by means of those theorems there will correspond an invariant point-set of the original

complex that can be found by means of the same theorems. This will be proved in section 6 and might be of use in determining invariant subsets of a given complex.

5. A large part of our method of finding invariant subsets consists in the application of the following composite statement of which we may suppress the proof:

THEOREM 5. *Invariant subsets, components, parts of certain local dimension, closures, sums, differences and intersections of invariant subsets of K^n are again invariant subsets of K^n.*

The rest of the method consists in repeated applications of generalizations of the following theorem proved in my dissertation: *

THEOREM 6. *The neighborhood complexes of corresponding points in homeomorphic complexes have the same homology characters; with the corollary: The set of points of K^n whose neighborhoods complexes have a definite set of homology characters is an invariant subset of K^n.*

The analysis of invariant subsets given by A. B. Brown † is founded on theorem 5 and a corollary of theorem 6: The subcomplex of K^n formed by all $(n-1)$-dimensional simplexes of K^n incident with exactly p $(p \neq 2)$ $n-1$ simplexes of K^n is invariant.

A first generalization of theorem 6 is the following:

THEOREM 7. *If in two homeomorphic complexes two corresponding points are interior points of p-simplexes, the complements of these p-simplexes have the same homology characters.*

Such points have as their neighborhood complexes the joins of the boundary of a p-simplex and the complements of their p-simplexes. Applying theorem 1 we find that the homology characters of those joins are obtained from the homology characters of the complements by adding p to the dimension of those characters, so that the first are invariants along with the second. From the character of the proof it is evident that this generalization does not increase our information on invariant subsets; however, it is interesting to compare theorems 7 and 10 and their proofs.

THEOREM 8. *Let M_1 and M_2 contained in M_1 be two invariant subsets of K^n and let L_1 and L_2 be the point-sets corresponding to M_1 and M_2 in the neighborhood complex of a point P of K^n, then the homology characters of L_1*

* *Loc. cit.*, p. 26.

† *American Journal of Mathematics*, Vol. 54 (1932), pp. 117-122.

modulo L_2 are invariants of K^n (and P). From this it follows: *The set of points where L_1 modulo L_2 has a definite set of homology characters is an invariant point-set of K^n.*

On the point-set K of K^n we consider as marked the following point-sets: P; the straight segments through P in K^n forming a point-set U; the sets N_1 and N_2 of those segments contained in M_1 and M_2; all point-sets corresponding to those in a complex $'K^n$ homeomorphic with K^n and marked with a prime. The intersections of U, N_1 and N_2 and a sufficiently small neighborhood of P are equal to the corresponding point-sets of $'K^n$. In a very small neighborhood of P we construct copies V and $'V$ of the neighborhood complexes of P in K^n and $'K^n$. The point-sets in those copies corresponding to the given invariant subsets will be called from now on: $L_1, L_2, 'L_1, 'L_2$. For any point of $U('U)$ except P we define a corresponding point on $V('V)$, called its projection on $V('V)$, defined as the point of intersection of the straight-segment in $K^n('K^n)$ through the given point and P with $V('V)$. To any singular chain of L_1 modulo L_2 we assign its projection on $'V$ which is a chain of $'L_1$ modulo $'L_2$. As boundaries correspond to boundaries under this projection we have defined a transformation of the group of homologies of L_1 modulo L_2 into that of $'L_1$ modulo $'L_2$. To complete the proof we must show: If $'\Gamma^p$ is the projection of a cycle Γ^p of L_1 modulo L_2 on $'V$ and if $''\Gamma^p$ is the projection of $'\Gamma^p$ on V then $\Gamma^p - ''\Gamma^p \sim 0$ on L_1 modulo L_2, for then the transformation of the two groups is isomorphic. $\Gamma^p - ''\Gamma^p$ is the boundary modulo M_2 of the chain formed by the straight segments in $'K^n$ between points of Γ^p and corresponding points of $'\Gamma^p$ plus the straight segments in K^n between points of $'\Gamma^p$ and corresponding points of $''\Gamma^p$. This chain is contained in U when V and $'V$ are taken sufficiently near P; its projection on V still has $\Gamma^p - ''\Gamma^p$ as its boundary but now taken modulo L_2, so that $\Gamma^p - ''\Gamma^p \sim 0$ on L_1 modulo L_2.

Theorem 8 can be generalized to give theorem 9 in the same way that we generalized theorem 6 to give theorem 7.

THEOREM 9. *If in two homeomorphic complexes two corresponding points are interior points of p-simplexes and $L_1(L_2)$ and $'L_1('L_2)$ are the sets in the complements of these p-simplexes corresponding to the invariant subset M_1 (M_2 contained in M_1) of one complex and the corresponding set of the other complex, then the homology characters of L_1 modulo L_2 are the same as those of $'L_1$ modulo $'L_2$.*

6. *If we find by means of theorems* (5) *and* (8) *an invariant point-set L in the complement U of a p-simplex S^p of K^n then the same theorems allow*

us to construct an invariant point-set in K^n of which the dimension is $p+1$ higher and to which L corresponds in the complement of S^p.

To prove this we must follow the process by which L was found and show that corresponding steps can be taken in K^n itself with corresponding results. This is obvious as far as theorem (5) is concerned. When we apply theorem (8) using the point P of U and point-sets L_1 and L_2 in the neighborhood complex V of P in U, the corresponding point-sets for K are on the join of P, S^p and V; corresponding to P we have to take an interior point Q of the join of P and S^p, the set corresponding to $L_1(L_2)$ in the neighborhood complex of Q is the join of $L_1(L_2)$ and the boundary R^p of the $(p+1)$-simplex $P \cdot S^p$. The homology characters of $R^p \cdot L_1 \bmod R^p \cdot L_2$ are, according to Section 3, obtained from those of $L_1 \bmod L_2$ by adding $p+1$ to the dimension of every character, so that the same sets of numbers are used to determine how the points Q of K^n and the points P of U can be divided into invariant point-sets.

As a result of theorem 6 the smallest invariant subsets we can find are open manifolds in the modern sense. The invariants of tensorial character defined for manifolds by J. W. Alexander (*Proc. Nat. Acad. Sci.* 1924) can be used to find still smaller invariant subsets and a connection between the results in a complex and in the complements of its simplexes as described just now does still exist. For a subsequent paper we plan a vindication of this statement and a closer analysis of these invariants.

7. The fundamental group can be used to split a complex into still smaller invariant subsets. However, owing to the fact that deformation invariants of dimensions larger than one have not been investigated closely the range of application is much smaller than that for instance of theorem 8, and the reasoning of section 6 cannot be extended in this case.

THEOREM 10. *If the interior of a p-simplex S^p of K^n is part of a p-dimensional invariant subset of K^n, then the fundamental group of (any component of) the point-set L of the complement U of S corresponding to any other invariant point-set M of K is a combinatorial invariant of K^n (and the two invariant subsets).*

This theorem can be used to distribute the p-simplexes of a p-dimensional invariant subset over smaller invariant subsets; the lower-dimensional simplexes to be distributed afterwards by means of theorem 5.

Because S^p is in a p-dimensional invariant subset of K^n we can find an open subset of S^p of which the corresponding subset in a complex $'K^n$ homeomorphic with K^n is open in a simplex $'S^p$ of $'K^n$. For any point of the join of $S^p('S^p)$ and its complement but not on $S^p('S^p)$ itself we define a corre-

sponding point on that complement called its projection on the complement and constructed by taking the point in the complement corresponding to the join of $S^p('S^p)$ and the given point. Let us represent the complements by sets on the point-set of our complex in a very small neighborhood of one of the points of the open set mentioned above so that they have a point of M in common. We chose that point in both the complements as the origin of the fundamental group of the point-sets corresponding in the complements to M. This causes the origins and so at the same time the elements of those fundamental groups to be transformed into each other under the projections defined before. Now we can repeat the substance of the proof of theorem 8, proving that L and the corresponding point-set constructed for $'K^n$ have isomorphic fundamental groups.

The restriction of our knowledge on deformation invariants explains the behavior of, for instance, the join of the boundary of a 2-simplex and a 3-dimensional manifold with a fundamental group different from the identity but whose 1-cycles are all homologous to zero. It seems very obvious that the boundary of the 2-simplex and the rest of the manifold are (minimal) invariant subsets of the join, but there does not seem to be an easy way of proving this.

8. A last theorem in this connection which we want to mention is the following:

THEOREM 11. *If in two homeomorphic n-dimensional complexes two corresponding points are both in $(n-p)$-simplexes, $p < 4$, the complements of those simplexes are homeomorphic.*

The proof is so simple that we shall not go into detail. For $p = 3$ the complements are divided into 2-dimensional manifolds by point-sets corresponding to $(n-1)$- and $(n-2)$-dimensional invariant subsets; the similarity of the way in which the manifolds have to be put together being decided by theorem 10 applied to those subsets $(p < 3)$ and the similarity of the corresponding manifolds being decided by means of theorem 9. The sign of the relation between any orientable manifold and a 1-dimensional invariant part of its boundary is not changed in passing to a homeomorphic complex because the sign of the relation between the chain formed by the manifold modulo all the rest of the boundary and the cycle formed by the chosen part of the boundary modulo its boundary is not changed in passing to a homeomorphic complex.

THE JOHNS HOPKINS UNIVERSITY.

SOME ADDITIONS TO THE THEORY OF COMBINATORS.*

By H. B. Curry.†

The present paper contains two amplifications of my dissertation *Grundlagen der kombinatorischen Logik.*‡ These are: (1) a revision of the definition of normal form (II D 1, Festsetzung 2) so that to every normal sequence there will correspond a unique normal combinator; (2) a generalization of the principal theorems on equality of combinators (II E 3, Sätze 4 and 5), so that the restriction that the combinators concerned be "eigentlich" is removed. These two parts of this paper are independent of one another, although both presuppose the thesis just mentioned.

Reference will be made to the dissertation by means of the subdivisions in that paper itself. Thus the above II D 1 means Kapitel II, Abschnitt D, § 1.

1. *Revised Definition of Normal Form.*

Convention 1. A regular combinator § whose terms are all of the form C_m, or $B_m I$) shall be called a *permutator*. (By II D 1, Festsetzung 4, permutators were denoted by $\mathfrak{C}$).

Convention 2. A permutator shall be called *normal*, or in the normal form, when it is either I or of the form

$$\mathfrak{C}_n \cdot \mathfrak{C}_{n-1} \cdot \mathfrak{C}_{n-2} \cdots \mathfrak{C}_2 \cdot \mathfrak{C}_1,$$

where each $\mathfrak{C}_k$ either is of the form

$$C_{a_k} \cdot C_{a_k+1} \cdot C_{a_k+2} \cdots C_{k-1}, \qquad (a_k < k),$$

or else, for $k \neq n$, $\mathfrak{C}_k \equiv I$.

Theorem 1. *For every permutator* $\mathfrak{C}$ *there exists one and only one normal permutator* $\mathfrak{C}'$ *such that*

$$\vdash \mathfrak{C} = \mathfrak{C}'. \tag{1}$$

Proof. If such a $\mathfrak{C}'$ exists it corresponds to a permutation (II C 4, Festsetzung 3) of order n (II D 6, Festsetzung 1). n must then also be the

* Presented to the American Mathematical Society, December 28, 1931.

† National Research Fellow.

‡ *American Journal of Mathematics*, Vol. 52 (1930), pp. 509-536, pp. 789-834.

§ II D 1, Festsetzung 1.

order of the permutation to which $\mathfrak{C}$ corresponds, (II D 6, Satz 3); hence n is uniquely determined.*

If U is any permutator let us denote by $U(k)$ the index of the variable which occupies the k-th place in the permutation to which U corresponds. Thus if $b_k = \mathfrak{C}(k)$ we have, if the x's be regarded as entities,

$$\vdash \mathfrak{C}x_0x_1x_2 \cdots x_n = x_0x_{b_1}x_{b_2} \cdots x_{b_n}.$$

I shall show that the numbers a_k determined as in Convention 2 by $\mathfrak{C}'$ are uniquely determined by the b_k, with the understanding that when $a_k = k$, then $\mathfrak{C}_k \equiv I$.

Suppose then a $\mathfrak{C}'$ is given, with the a_k defined in terms of $\mathfrak{C}'$ as in Convention 2 with the understanding just mentioned. Then

$$\begin{aligned} \mathfrak{C}_k(j) &= j, \quad \text{if} \quad j > k; \\ \mathfrak{C}_k(k) &= a_k. \end{aligned} \tag{2}$$

Now it is easily verified that for permutators U and V

$$(U \cdot V)(k) = U(V(k));$$

hence

$$(\mathfrak{C}_k \cdot \mathfrak{C}_{k-1} \cdots \mathfrak{C}_1)(k) = a_k. \tag{3}$$

Consequently, if we define temporarily

$$D_k \equiv \mathfrak{C}_n \cdot \mathfrak{C}_{n-1} \cdots \mathfrak{C}_{k+1}, \qquad D_n \equiv I, \tag{4}$$

so that

$$\vdash \mathfrak{C}' = D_k \cdot (\mathfrak{C}_k \cdot \mathfrak{C}_{k-1} \cdots \mathfrak{C}_1),$$

then, by (2), (3),

$$\mathfrak{C}'(k) = D_k(a_k) = D_{k+1}(k). \tag{5}$$

If $\mathfrak{C}'$ is to satisfy the conditions, $\mathfrak{C}'(k)$ must equal b_k, hence

$$b_k = D_k(a_k),$$

or, since permutators have unique universes,

$$a_k = D_k^{-1}(b_k). \tag{6}$$

The equations (6) determine the a's uniquely. In fact I shall prove that for any $r < n$ there exists a unique set of numbers $a_n, a_{n-1}, \cdots, a_{n-r}$ such that: (1) $a_n < n$, and for $k \leqq r$, $a_{n-k} \leqq n - k$; (2) if we define $\mathfrak{C}_{n-k}$ as in Convention 2 for $a_{n-k} < n - k$, and as I for $a_{n-k} = n - k$, then the last $r + 1$

* These statements are true even when $\mathfrak{C}' \equiv I$, $n = 0$.

equations (6) are satisfied, D_k being defined by (4). This assertion is true for $r = 0$; for the last equation (6) is

$$a_n = b_n,$$

which determines a_n uniquely; moreover $a_n < n$ since $\mathfrak{C}$ corresponds to a sequence of order n. Suppose now the assertion is true for a given r. Then the last $r + 1$ equations (6) determine $a_n, a_{n-1}, \cdots, a_{n-r}$ uniquely and subject to the inequalities stated. The next preceding equation, viz.,

$$a_{n-r-1} = D^{-1}{}_{n-r-1}(b_{n-r-1}),$$

determines a_{n-r-1} uniquely, since $D^{-1}{}_{n-r-1}$ contains no number not already determined, and $a_{n-r-1} \leqq n$. Suppose $a_{n-r-1} = j > n - r - 1$. Then

$$\begin{aligned} D_{n-r-1}(j) &= (D_j \cdot \mathfrak{C}_j \cdot \mathfrak{C}_{j-1} \cdots \mathfrak{C}_{n-r})(j) && \text{(by def. of } D_{n-r-1}), \\ &= (D_j \cdot \mathfrak{C}_j)(j) && \text{(by (2))}, \\ &= D_j(a_j) = b_j && \text{(by (5))}. \end{aligned}$$

But,

$$D_{n-r-1}(j) = b_{n-r-1} \qquad \text{(by (6))}. \tag{7}$$

$\therefore b_j = b_{n-r-1}$, which is impossible if $j > n - r - 1$ because the b's are all distinct. Therefore $a_{n-r-1} \leqq n - r - 1$, and the above assertion, if true for r, is also true for $r + 1$. By induction it is therefore true for every $r < n$; hence in particular for $r = n - 1$. The equations (6) have therefore one and only one set of solutions $a_1, a_2, \cdots, a_n$.

Now let the $a_1, a_2, \cdots, a_n$ be determined as in the last paragraph, and let $\mathfrak{C}'$, D_k, be defined in terms of them. Then by (5), which is still true, $\mathfrak{C}'$ satisfies the conditions of the theorem, and is by what has just been proved the only normal permutator which does so.

Convention 3. A regular combinator shall be said to be *normal*, or in the normal form, when (except for parentheses and the omission of identical terms) it is in the form

$$\mathfrak{K} \cdot \mathfrak{W} \cdot \mathfrak{C} \cdot \mathfrak{B}\ *$$

(1) $\mathfrak{K}$, $\mathfrak{W}$, $\mathfrak{B}$, are in the normal forms already defined (II D 1, Festsetzung 2);

(2) $\mathfrak{C}$ is in the form of Convention 2;

(3) $\mathfrak{C}$ corresponds to the permutation π determined as follows: let η be

* Any of the components $\mathfrak{K}$, $\mathfrak{W}$, $\mathfrak{C}$, $\mathfrak{B}$, may reduce to the identical combinator, I, and hence be omitted.

the sequence (Folge) to which the given combinator corresponds; determine from η a γ and an ω as in II C 5, Satz 1; from ω a κ and a μ as in II C 4, Satz 1; and from μ an α satisfying the conditions for the first factor in II C 4, Satz 3; then of those permutations which satisfy the conditions for the second factor of the last theorem choose for π that one which does not permute among themselves any of the letters $x_{c_1}, x_{c_2}, \cdots x_{c_p}$, which occupy places held by the same letters in α.*

THEOREM 2. *To every normal sequence there corresponds one and only one normal combinator.*

Proof. Let η be the given sequence, and let γ, ω, κ, μ, α, π be determined by the process outlined in Convention 3. Then

$$\eta = \kappa \cdot \alpha \cdot \pi \cdot \gamma.$$

Let $\mathfrak{K}$, $\mathfrak{W}$, $\mathfrak{C}$, $\mathfrak{B}$ be the normal combinators which correspond to κ, α, π, and γ, respectively (II C 4, Sätze 2 and 4, Theorem 1 above, and II C 3, Satz 3). Then $(\mathfrak{K} \cdot \mathfrak{W} \cdot \mathfrak{C} \cdot \mathfrak{B})$ is normal and corresponds to η.

Suppose there were another such combinator; let it be $(\mathfrak{K}' \cdot \mathfrak{W}' \cdot \mathfrak{C}' \cdot \mathfrak{B}')$. By II C 5, Satz 2, these combinators can differ only in their $\mathfrak{C}$'s. By the third requirement in Convention 3, both $\mathfrak{C}$ and $\mathfrak{C}'$ must correspond to π as described in that requirement. By the theorems there mentioned, ω, μ, α, and therefore π, are uniquely determined by η.† Hence $\mathfrak{C}$ and $\mathfrak{C}'$ must both correspond to the same permutation, and therefore, by Theorem 1, are identical.

THEOREM 3. *If $\mathfrak{R}$ is a regular combinator, then there exists one and only one normal combinator $\mathfrak{R}'$ such that*

$$\vdash \mathfrak{R} = \mathfrak{R}'.$$

Proof. Follows at once by II D 6, Sätze 2, 3, 4, and Theorem 2.

2. *General Theorem on Equality of Combinators.*

Convention 4. If $\mathfrak{X}$ and $\mathfrak{Y}$ are expressions involving the letters $x_1, x_2, \cdots, x_m$; then

$$\vdash_* \mathfrak{X} = \mathfrak{Y}$$

* This π is the same as the β constructed as an example in my paper "An Analysis of Logical Substitution," *American Journal of Mathematics*, Vol. 51 (1929), pp. 363-384, Lemma 4; it being understood that the γ of that paper is the same as the μ mentioned here. The other permutations which satisfy the conditions for the second factor in II C 4, Satz 3 are the β's discussed in Lemma 5 of the earlier paper, from which it appears that π is uniquely described.

† See footnote above.

shall mean that if the $x_1, x_2, \cdots x_m$ be treated as variables,—i. e. as denoting indeterminate entities,—then the equality of $\mathfrak{X}$ and $\mathfrak{Y}$ can be proved. Similarly if $\mathfrak{X}$ and $\mathfrak{Y}$ also contain the letters $y_1, y_2, \cdots, y_n$, $\vdash_{x,y} \mathfrak{X} = \mathfrak{Y}$ shall mean that the equality can be proved if the x's and y's be treated as variables. (These conventions apply even when the x's and y's actually denote entities. Cf. II E 3, Festsetzung 3).

Convention 5. A formula, such as $\vdash X = Y$, shall be said to *follow combinatorially* when it can be derived, as in II, from the combinatory axioms, the combinatory rules B, C, W, K, and the properties of equality proved in $I\ D$. This may be symbolized thus

$$\vdash X = Y \qquad \text{(combinatorially)}.$$

The same shall apply to formulas of type considered in Convention 4.

THEOREM 4. *If $\mathfrak{X}$ and $\mathfrak{Y}$ are combinations of variables $x_1, x_2, \cdots, x_n$, and entities $u_1, u_2, \cdots, u_m$,* and X and Y are combinators such that,—*

(i) *the expressions* $(XIu_1u_2 \cdots u_m\, x_1x_2 \cdots x_n)$ *and* $(YIu_1u_2 \cdots u_m\, x_1x_2 \cdots x_n)$ *reduce formally to $\mathfrak{X}$ and $\mathfrak{Y}$ respectively;*

(ii) *it follows combinatorially that*

$$(1) \qquad \vdash_x \mathfrak{X} = \mathfrak{Y};$$

then it follows combinatorially that

$$(2) \qquad \vdash XIu_1u_2 \cdots u_m = YIu_1u_2 \cdots u_m.$$

Proof. Suppose first that $\mathfrak{X}$ differs from $\mathfrak{Y}$ only in that a certain expression $\mathfrak{A}$ in $\mathfrak{X}$ is replaced in $\mathfrak{Y}$ by $\mathfrak{B}$, where

$$(3) \qquad \vdash_x \mathfrak{A} = \mathfrak{B} \qquad \text{(combinatorially)}.$$

Then there must exist a regular combinator T such that the expressions $(TI\mathfrak{A}u_1u_2 \cdots u_m\, x_1x_2 \cdots x_n)$ and $(TI\mathfrak{B}u_1u_2 \cdots u_m\, x_1x_2 \cdots x_n)$ reduce formally to $\mathfrak{X}$ and $\mathfrak{Y}$ respectively.† We shall consider two cases.

Case 1. Suppose $\mathfrak{A}$ and $\mathfrak{B}$ are entities, and that the formula (3) is established without the use of variables. Let U and V be combinators such

* It is of course not required that all these symbols appear in both $\mathfrak{X}$ and $\mathfrak{Y}$.

† For this reduction $\mathfrak{X}$ is to be regarded as a combination of the symbols $\mathfrak{A}, u_1, u_2, \cdots, u_m, x_1, x_2, \cdots, x_n$, and $\mathfrak{A}$ appears only once in $\mathfrak{X}$. Similarly for the relation between $\mathfrak{Y}$ and $\mathfrak{B}$. Superfluous u's can be cancelled by factors K_j appearing in T.

that $UIu_1u_2 \cdots u_m$ and $VIu_1u_2 \cdots u_m$ formally reduce respectively to $\mathfrak{A}$ and $\mathfrak{B}$, and let S be a regular combinator corresponding to the sequence

$$x_0x_2(x_1x_2x_3x_4 \cdots x_{m+2})x_3x_4 \cdots x_{m+2}x_{m+3} \cdots .$$

Let $X' \equiv STU$, $Y' \equiv STV$. Then the expressions $(X'Iu_1u_2 \cdots u_m\, x_1x_2 \cdots x_n)$ and $(Y'Iu_1u_2 \cdots u_m\, x_1x_2 \cdots x_n)$ reduce formally to $\mathfrak{X}$ and $\mathfrak{Y}$ respectively.* Therefore, by II E 3, Satz 4, $\vdash XI = X'I$ and $\vdash YI = Y'I$, because, for example, if the u's be regarded as variables the two expressions

$$XIu_1u_2 \cdots u_m\, x_1x_2 \cdots x_n \quad \text{and} \quad X'Iu_1u_2 \cdots u_m\, x_1x_2 \cdots x_n$$

both reduce to the same sequence of variables and hence XI and $X'I$ correspond to the same sequence of variables. On the other hand

$$\vdash X'Iu_1u_2 \cdots u_m = TI\mathfrak{A}u_1u_2 \cdots u_m,$$
$$\vdash Y'Iu_1u_2 \cdots u_m = TI\mathfrak{B}u_1u_2 \cdots u_m.$$

Since the two left-hand sides reduce formally to their respective right-hand sides. The two right-hand sides are equal by (3). Hence, finally,

$$\begin{aligned} \vdash XIu_1u_2 \cdots u_m &= X'Iu_1u_2 \cdots u_m \\ &= Y'Iu_1u_2 \cdots u_m \\ &= YIu_1u_2 \cdots u_m. \end{aligned}$$

which proves the theorem for this case.

Case 2. Suppose $\mathfrak{A}$ differs from $\mathfrak{B}$ in that $\mathfrak{A}$ reduces to $\mathfrak{B}$ by application of a single one of the combinatory rules, B, C, W, K. (There is no loss of generality in supposing that $\mathfrak{A}$ reduces to $\mathfrak{B}$). Let U be the combinator concerned in this reduction; it is then one of the u_i. Let $\mathfrak{X}'$ be the expression obtained from $\mathfrak{X}$ by replacing U as it appears at the beginning of $\mathfrak{A}$ by a new variable y, but otherwise leaving $\mathfrak{X}$ unchanged.† Let X' be the regular combinator such that $X'Iyu_1u_2 \cdots u_m\, x_1x_2 \cdots x_n$ reduces formally to $\mathfrak{X}'$. Let V be the regular combinator such that

$$\vdash_{x,u} Vx_0u_1u_2 \cdots u_m = x_0Uu_1u_2 \cdots u_m,$$

(such a V exists since U is simply one of the u_i). Then the expression

* The first one reduces formally to $TI(UIu_1u_2 \cdots u_m)u_1u_2 \cdots u_m\, x_1x_2 \cdots x_n$, thence to $\mathfrak{X}$ with $\mathfrak{A}$ replaced by $UIu_1u_2 \cdots u_m$. In this the only symbol not considered formally is U, so that the reduction proceeds formally to $\mathfrak{X}$. Similarly for the second one.

† If U appears in $\mathfrak{X}$ in more than one place, this replacement is to be made only at the beginning of $\mathfrak{A}$.

$(V \cdot X')Iu_1u_2 \cdots u_m x_1x_2 \cdots x_n$ reduces formally to $\mathfrak{X}$ (the u_i, x_j all being regarded as variables). Hence by II E 3, Satz 4, (cf. above)

$$(4) \qquad \vdash (V \cdot X')I = XI.$$

On the other hand if we regard U as a combinator and $u_1, \cdots, u_m, x_1, \cdots x_n$, as variables, the expression $X'IUu_1u_2 \cdots u_p x_1x_2 \cdots x_n$ reduces to $\mathfrak{Y}$. Hence, (II E 3, Satz 4)

$$(5) \qquad \vdash X'IU = YI.$$

Therefore

$$\begin{aligned} \vdash XIu_1u_2 \cdots u_m &= (V \cdot X')Iu_1u_2 \cdots u_m && \text{(By (4))}, \\ &= V(X'I)u_1u_2 \cdots u_m && \text{(II B 4, Satz 1)}, \\ &= X'IUu_1u_2 \cdots u_m && \text{(Def. of } V\text{)}, \\ &= YIu_1u_2 \cdots u_m && \text{(By (5))}. \end{aligned}$$

The theorem is therefore proved in case the transition from $\mathfrak{X}$ to $\mathfrak{Y}$ is effected by means of a single formula of type (3). For if (3) be an axiom or a previously proved formula we have the Case 1; if (3) be a special case of a rule, Case 2.

In the most general case there must exist a series of expressions $\mathfrak{X}_1, \mathfrak{X}_2, \cdots, \mathfrak{X}_p$, where $\mathfrak{X}_1 \equiv \mathfrak{X}$, $\mathfrak{X}_p \equiv \mathfrak{Y}$, such that $\mathfrak{X}_{i+1}$ is derived from $\mathfrak{X}_i$ by a single formula (3) in the manner just considered. If X_i be a combinator related to $\mathfrak{X}_i$ in the same manner as X is related to $\mathfrak{X}$ and Y to $\mathfrak{Y}$,* then by the preceding proof,

$$\vdash X_i u_1u_2 \cdots u_m = X_{i+1}u_1u_2 \cdots u_m.$$

Therefore, since equality is transitive,

$$\vdash Xu_1u_2 \cdots u_m = Yu_1u_2 \cdots u_m. \qquad \text{q. e. d.}\dagger$$

---

* Such combinations of course exist; in fact we can find an X_i of the form $\mathfrak{N}_i I$, where $\mathfrak{N}_i$ is normal.

† In case the $\mathfrak{X}_i$ contain letters appearing in neither $\mathfrak{X}$ nor $\mathfrak{Y}$ the following modification of the proof is required. Let $v_1, v_2, \cdots, v_p$, $y_1, y_2, \cdots, y_q$ be the extra letters so appearing, the v's representing entities and the y's being variables. Let

$$V_r = K_r \cdot K_{r-1} \cdots K_2 \cdot K_1, \qquad r = p + q.$$

Then by the method in the text we can establish the formula

$$\vdash (B_m V_r \cdot X) I u_1 u_2 \cdots u_m v_1 v_2 \cdots v_p = (B_m V_r \cdot Y) I u_1 u_2 \cdots u_m v_1 v_2 \cdots v_p,$$

for if we insert first the y's and then the x's after the two sides the resulting expressions reduce formally to $\mathfrak{X}$ and $\mathfrak{Y}$ respectively. Hence if $w_1, w_2, \cdots, w_q$ are any entities (say $w_i \equiv I$), we have

$$\begin{aligned} \vdash (B_m V_r \cdot X) I u_1 u_2 \cdots u_m v_1 \cdots v_p w_1 \cdots w_q \\ = (B_m V_r \cdot Y) I u_1 u_2 \cdots u_m v_1 \cdots v_p w_1 \cdots w_q, \end{aligned}$$

whence, by reduction of both sides, we have (2).

Remark. That the above theorem is more general than those of II E 3, is shown by the following example, which may be proved by its aid,

$$\vdash C(W \cdot BB)I = WCI.$$

Neither of the combinators is "eigentlich"; in fact if $x_0 x_1$ be written after each they both reduce to $x_0 I x_1$. On the other hand if I is to be regarded formally the right-hand side reduces to the expression just written, while the reduction of the left-hand side cannot get any further than $x_0 I(Ix_1)$.

THEOREM 5. *If X any Y are entities such that it follows combinatorially that*

$$\vdash_x Xx_1x_2 \cdots x_n = Yx_1x_2 \cdots x_n;$$

then it follows combinatorially that

$$\vdash X = Y.$$

Proof. Let X' and Y' be combinators, and $u_1, u_2, \cdots u_m$ entities such that $X'Iu_1u_2 \cdots u_m$ and $Y'Iu_1u_2 \cdots u_m$ reduce formally to X and Y respectively. Then

$$\begin{aligned} \vdash X &= X'Iu_1u_2 \cdots u_m && \text{(by const.)}, \\ &= Y'Iu_1u_2 \cdots u_m && \text{(Theorem 4)}, \\ &= Y \quad \text{q. e. d.} && \text{(by const.)}. \end{aligned}$$

UNIVERSITY OF CHICAGO.

ON THE OVERCONVERGENCE OF SEQUENCES OF RATIONAL FUNCTIONS.*

By J. L. WALSH.

If a sequence of rational functions of the complex variable z converges to a given function $f(z)$ on a given point set C with a certain degree of approximation, then that sequence may necessarily converge to the given function or its analytic extension on a larger point set containing C in its interior. For instance, it was shown by S. Bernstein that convergence of the sequence of polynomials $p_n(z)$ of respective degrees $n = 1, 2, \cdots$ in such a way that we have for some $R > 1$

$$| f(z) - p_n(z) | \leqq M/R^n, \qquad (-1 \leqq z \leqq 1),$$

implies uniform convergence of the sequence $p_n(z)$ in some ellipse whose foci are the points 1 and -1. Various generalizations of this result have been studied by the present writer, where the degree of approximation of polynomials is given in a certain way not on the point set $-1 \leqq z \leqq 1$, but on a more general point set,† and also where the approximating functions are rational functions more general than polynomials and the point set is a region or rectifiable curve.‡ It is the object of the present paper to set forth a result which includes all of those mentioned, and which indeed seems to be the most general result of its type. In the previous study, we considered *primarily* approximation in a circular region by rational functions whose poles lie in another circular region; in the present paper more general point sets instead of circular regions are contemplated in these two rôles. We shall first prove an interesting special case, and then indicate the modifications for the more general theorem.

I.

Let C be an arbitrary closed Jordan region of the extended § z-plane and let D be an arbitrary closed point set having no point in common with C.

* Presented to the Society, March 25, 1932.

† *Münchner Berichte* (1926), pp. 223-229.

‡ *Transactions of the American Mathematical Society*, Vol. 30 (1928), pp. 838-847 and Vol. 34 (1932), pp. 22-74.

§ That is to say, we adjoin the point at infinity to the usual (finite) z-plane; consequently the notions *closure* (of a point set), *Jordan curve*, and *Jordan region* are most conveniently interpreted on the sphere (stereographic projection of the plane) instead of in the plane itself.

Let $C_R(z')$ denote the curve in the z-plane which is the image of the circle $|w| = R > 1$ when the complement of C is mapped onto $|w| > 1$ in such a way that the point $z = z'$ of D corresponds to $w = \infty$. The closed regions containing C bounded by the various curves $C_R(z')$ have obviously the region C in common when z' varies over D, and conceivably have a still larger point set in common. We denote by C_R this common point set whatever it may be. It will appear later that the point set common to the regions considered is itself closed, and it will also appear that C actually lies interior to C_R.

LEMMA I. *If $r_n(z)$ is a rational function of z of degree n* whose poles lie in D and if we have*

(1) $$|r_n(z)| \leqq M, \quad \text{for } z \text{ on } C,$$

then we have also

(2) $$|r_n(z)| \leqq MR^n, \quad \text{for } z \text{ on } C_R.$$

Let $z_1, z_2, \cdots, z_n$ be the poles of $r_n(z)$; these lie in D and need not be distinct. The function $r_n(z)$ need not have n poles; if there are fewer than n poles, only an obvious modification is to be made in our reasoning. Let $w = \phi_i(z)$ be a function which maps the exterior of C onto the exterior of $\gamma: |w| = 1$ so that the point $z = z_i$ corresponds to $w = \infty$. The function $\phi_i(z)$ is analytic exterior to C and (if suitably defined on C itself) continuous in the corresponding closed region except for a pole of the first order at $z = z_i$. The modulus of $\phi_i(z)$ on C is unity. The function

$$r_n(z)/\phi_1(z)\phi_2(z) \cdots \phi_n(z)$$

is analytic exterior to C, continuous in the corresponding closed region, and for z on C we have by (1)

(3) $$|r_n(z)/\phi_1(z)\phi_2(z) \cdots \phi_n(z)| \leqq M.$$

Since inequality (3) holds for z on C it holds also for z exterior to C. Throughout C_R exterior to C we have

$$|\phi_i(z)| \leqq R,$$

so that (2) now follows from (3) for z in C_R exterior to C. But C_R contains C in its interior, and $r_n(z)$ has no poles in C_R, so (2) is valid for z anywhere in C_R.

---

* That is to say, $r_n(z)$ can be written in the form

$$(a_0z^n + a_1z^{n-1} + \cdots + a_n)/(b_0z^n + b_1z^{n-1} + \cdots + b_n),$$

where the denominator does not vanish identically.

Lemma I is analogous to, and a generalization of, a lemma used by Bernstein in the proof of his result already mentioned. The present method is closely related to that used by M. Riesz * in his proof of Bernstein's lemma.

Lemma I can obviously be extended to apply to point sets C more general than Jordan regions; we shall discuss this fact in more detail later. Let us return to the original case considered and use our original notation.

LEMMA II. *If C is a Jordan region and if the point set D (having no point in common with C) is a Jordan region bounded by a curve which can be denoted by $C_\rho(\zeta)$ [where ζ is some point of D], then the point set C_R is the Jordan region containing C bounded by the curve $C_\nu(\zeta)$, where*

$$\nu = (R\rho + 1)/(R + \rho).$$

Let $w = \phi(z)$ be a function which maps the exterior of C onto the exterior of $\gamma: |w| = 1$ so that the point $z = \zeta$ corresponds to $w = \infty$. Then the function

$$w' = (1 - \bar{w}_1 w)/(w - w_1) = [1 - \overline{\phi(z_1)}\phi(z)]/[\phi(z) - \phi(z_1)]$$

maps the exterior of C onto the exterior of $\gamma': |w'| = 1$ so that $z = z_1$ corresponds to $w' = \infty$. The restrictions $|w_1| \geqq \rho$, $|\phi(z_1)| \geqq \rho$, z_1 in the closed region not containing C bounded by $C_\rho(\zeta)$, z_1 in D, are all equivalent. All points z not in C_R can be expressed by the conditions

$$\left| \frac{1 - \overline{\phi(z_1)}\phi(z)}{\phi(z) - \phi(z_1)} \right| \geqq R, \qquad |\phi(z_1)| \geqq \rho, \tag{4}$$

or by the equivalent conditions

$$\left| \frac{1 - \bar{w}_1 w}{w - w_1} \right| \geqq R, \qquad |w_1| \geqq \rho, \tag{5}$$

where $w = \phi(z)$, $w_1 = \phi(z_1)$.

Conditions (5), considered as restrictions on w, are not difficult to transform. We substitute

$$w' = (1 - \bar{w}_1 w)/(w - w_1),$$

and write (5) in the form

$$w = (w_1 w' + 1)/(w' + \bar{w}_1), \qquad |w'| \geqq R > 1, \qquad |w_1| \geqq \rho > 1. \tag{6}$$

Here w' and w_1 are arbitrary, subject to the conditions indicated, and we seek the locus of w. The transformation in (6) transforms the unit circle

---

* *Acta Mathematica*, Vol. 40 (1916), pp. 337-347.

$|w'|=1$ and its exterior into $|w|=1$ and its exterior, so we obviously have $|w|>1$. Let us set $|w'|\geqq R_1\geqq R$, $|w_1|=\rho_1\geqq\rho$, so we have by (6)

$$w\bar{w}=(R_1^2\rho_1^2+w_1w'+\bar{w}_1\bar{w}'+1)/(R_1^2+w_1w'+\bar{w}_1\bar{w}'+\rho_1^2).$$

This last expression can be written as

$$\frac{R_1^2\rho_1^2+2x+1}{R_1^2+2x+\rho_1^2}, \tag{7}$$

a fraction whose numerator is greater than its denominator, which is positive. We are considering the fraction for $-\rho_1R_1\leqq x\leqq\rho_1R_1$. The denominator vanishes for $x=-(R_1^2+\rho_1^2)/2\leqq-R\rho$, so, since the function (7) is represented by an equilateral hyperbola whose asymptotes are the axes, its minimum occurs for $x=\rho_1R_1$, and this minimum value is

$$[(R_1\rho_1+1)/(R_1+\rho_1)]^2.$$

The relation

$$(R_1\rho_1+1)/(R_1+\rho_1)\geqq(R\rho+1)/(R+\rho)$$

is a consequence of the conditions $R_1\geqq R$, $\rho_1\geqq\rho$, so we have shown that *the complement of C_R lies in the Jordan region not containing C bounded by the curve $C_\nu(\zeta)$, $\nu=(R\rho+1)/(R+\rho)$.*

Reciprocally, let z be given in the Jordan region not containing C bounded by this curve $C_\nu(\zeta)$; we shall prove that z belongs to the complement of C_R. When this situation is transformed onto the w-plane, it reduces to the following. Given w, in absolute value not less than $(R\rho+1)/(R+\rho)$; to find w' and w_1 such that we have (6) satisfied.

If $|w|\geqq R$, it suffices to set $w'=w$, $w_1=\infty$. If we have

$$R>|w|\geqq(R\rho+1)/(R+\rho),$$

we define the quantity ρ_1 by the equation

$$|w|=(R\rho_1+1)/(R+\rho_1)$$

so we have $\rho_1\geqq\rho$. Then (6) is satisfied if we merely set

$$w_1=\rho_1w(R+\rho_1)/(R\rho_1+1),\qquad w'=R(R\rho_1+1)/w(R+\rho_1);$$

we have $|w_1|=\rho_1$, $|w'|=R$, and Lemma II is completely proved.

It follows from Lemma II that the complement of C_R never has a point in common with C, no matter how D (having no point in common with C) may be chosen. For such an arbitrary D lies in some Jordan region D' bounded by a Jordan curve $C_\rho(\zeta)$. Enlargement of the original D to the new D' merely adds points to the complement of C_R or leaves that complement

unchanged, and the complement of the C_R corresponding to D' has no point in common with C. Thus C itself always lies interior to C_R.

It is also true under our present hypothesis, namely that C is a Jordan region and D an arbitrary closed set having no point in common with C, that the point set C_R is closed. In fact, C_R is the point set common to a finite or infinite number of closed regions (each bounded by a curve $C_R(z')$ and containing C) and therefore closed.

It is not true under our present hypothesis that C_R is necessarily connected, as we proceed to illustrate by an example. Let us choose C as the unit circle, and D as the point set 2, 2ω, $2\omega^2$, where ω is an imaginary cube root of unity. The curve $C_R(z')$ is a circle of the coaxial family determined by z' (considered as a null circle) and C. For large values of R, the three circles $C_R(2)$, $C_R(2\omega)$, $C_R(2\omega^2)$ are small circles which do not intersect each other and which contain the points 2, 2ω, $2\omega^2$ respectively in their interiors. As R decreases, those circles vary as follows: they increase in size, intersect each other, become straight lines, become circles not containing the respective points 2, 2ω, $2\omega^2$ but containing C, decreases in size, and finally approach C. At the stage where these circles $C_R(z')$ are still circles containing the respective points 2, 2ω, $2\omega^2$ but intersecting each other, the point at infinity is exterior to the circles $C_R(z')$ and hence a point of C_R, yet cannot be connected with C by a broken line consisting wholly of point of C_R.

By means of Lemma I we shall be able to prove

THEOREM I. *Let $r_n(z)$ be a sequence of rational functions of respective degrees n such that we have for z in an arbitrary closed Jordan region C*

$$(8) \qquad |f(z) - r_n(z)| \leqq M/\rho^n, \qquad \rho > 1.$$

If the poles of the functions $r_n(z)$ lie in the closed set D, where D has no point in common with C, then the sequence $r_n(z)$ converges uniformly for z in C_R, where $R < \rho^{1/2}$.

If (8) is valid for z in C and if the function $r_{n+1}(z) - r_n(z)$ is of degree $n+1$, with its poles in D, then the sequence $r_n(z)$ converges uniformly for z in C_R, where $R < \rho$.

If (8) is valid for z in C and if there exists a sequence $r'_n(z)$ of functions such that $r'_n(z) - r_n(z)$ is a rational function of degree n with its poles in D, if we have

$$(9) \qquad |f(z) - r'_n(z)| \leqq M/\rho^n, \qquad z \text{ in } C,$$

and if the sequence $r'_n(z)$ converges uniformly for z in C_R, where $R < \rho$ (R may or may not depend on the sequence $r'_n(z)$), then the sequence $r_n(z)$ also converges uniformly for z in C_R.

We prove the various parts of Theorem I in order. From the two inequalities

$$| f(z) - r_n(z) | \leqq M/\rho^n, \qquad | f(z) - r_{n+1}(z) | \leqq M/\rho^{n+1},$$

satisfied for z on C, we have

$$| r_{n+1}(z) - r_n(z) | \leqq \frac{M}{\rho^n}\left(1 + \frac{1}{\rho}\right),$$

also for z on C. The rational function $r_{n+1}(z) - r_n(z)$ is of degree $2n+1$, so Lemma I yields for z on C_R,

$$| r_{n+1}(z) - r_n(z) | \leqq \frac{M}{\rho^n}\left(1 + \frac{1}{\rho}\right) R^{2n+1} = MR\left(1 + \frac{1}{\rho}\right)\left(\frac{R^2}{\rho}\right)^n.$$

Hence the sequence $r_n(z)$ converges as we have asserted.

It follows also from the proof just given that if $r_{n+1}(z) - r_n(z)$ is of degree $n+1$, with its poles in D, we have for z on C_R

$$| r_{n+1}(z) - r_n(z) | \leqq \frac{M}{\rho^n}\left(1 + \frac{1}{\rho}\right) R^{n+1},$$

so the sequence $r_n(z)$ converges uniformly for z on C_R, $R < \rho$. The assumption that $r_{n+1}(z) - r_n(z)$ is a rational function of degree $n+1$ is satisfied for instance by the sequence of functions corresponding to a series of the type*

$$a_0 + a_1 \frac{z - \beta_1}{z - \alpha_1} + a_2 \frac{(z-\beta_1)(z-\beta_2)}{(z-\alpha_1)(z-\alpha_2)} + a_3 \frac{(z-\beta_1)(z-\beta_2)(z-\beta_3)}{(z-\alpha_1)(z-\alpha_2)(z-\alpha_3)} + \cdots$$

If we have both (8) and (9) valid, we obtain similarly for z on C,

$$| r_n(z) - r'_n(z) | \leqq 2M/\rho^n,$$

and Lemma I gives the inequality

$$| r_n(z) - r'_n(z) | \leqq (2M/\rho^n) R^n$$

for z on C_R. Thus the sequence $r_n(z) - r'_n(z)$ converges uniformly for z on C_R, $R < \rho$. The sequence $r'_n(z)$ is known to converge uniformly for z on C_R, so it follows that the sequence $r_n(z)$ converges also in this manner, and the proof of Theorem I is complete. If (9) is used, we do not need to know that $r_n(z)$ is rational or of degree n.

In each of the parts of Theorem I it is naturally true that the sequence $r_n(z)$ converges to the function $f(z)$ (or its analytic extension) for z in C_R if C_R is connected and otherwise for z in that connected part of C_R containing C; for by (8) itself the limit of the sequence is $f(z)$ on C, and the limit of the sequence is analytic for z on C_R.

* Compare Angelescu, *Bulletin de l'Académie Roumaine*, Vol. 9 (1925), pp. 164-168; Walsh, *Proceedings of the National Academy of Sciences*, Vol. 18 (1932), pp. 165-171.

II.

It is clear that the reasoning used in proving Lemmas I and II can be extended to include point sets C much more general than Jordan regions; the reasoning holds with only minor changes if C is any region or other closed point set whose complement is connected. But if the complement of C is not connected, our previous discussion requires to be supplemented, and we proceed to indicate how that can be accomplished.

Let C and D be arbitrary closed point sets with no common point, such that the Dirichlet problem (for arbitrary boundary values) has a solution for every region $S(z')$ defined as the totality of points which can be joined to a given point z' of D by broken lines lines not meeting C. Such a totality of points actually forms an open region and its boundary consists entirely of points of C, as the reader will easily prove.* As before, let $C_R(z')$ denote the curve or curves in the z-plane which are the image of the circle $|w| = R < 1$ when the region $S(z')$ is mapped onto $|w| > 1$ in such a way that the point $z = z'$ corresponds to $w = \infty$. The mapping need not be smooth and the mapping function $w = \phi(z)$ is not uniquely defined, but the mapping function exists by the hypothesis on the region $S(z')$, and the point set $C_R(z')$ is uniquely defined. Let $S_R(z')$ denote the complement in the z-plane of the set $|\phi(z)| > R > 1$, so that $S_R(z')$ is closed and contains C, and is bounded by $C_R(z')$. Let C_R denote the point set common to all the point sets $S_R(z')$, where z' takes on all possible positions in D; the given set C surely belongs to C_R. It can be proved as in connection with Lemma II that C lies *interior* to C_R, and that C_R is closed.

LEMMA III. *If $r_n(z)$ is a rational function of z of degree n whose poles lie in D and if we have*

$$(1) \qquad |r_n(z)| \leqq M, \qquad z \text{ on } C,$$

then we have also

$$(2) \qquad |r_n(z)| \leqq MR^n, \qquad z \text{ on } C_R.$$

The proof of Lemma III is slightly more complicated than that of Lemma I because of the fact that in the present case the point set C may separate the plane. Let $z_1, z_2, \cdots, z_n$ be the poles of $r_n(z)$, not necessarily all distinct; if there are fewer than n poles, obvious modifications are to be made in our discussion. Let $z_1, z_2, \cdots, z_m$ be the poles of $r_n(z)$ which lie in $S(z_1)$. Denote by $\phi_1(z), \phi_2(z), \cdots, \phi_m(z)$ functions which map the region $S(z_1)$ onto the exterior of $\gamma: |w| = 1$ so that the respective points $z_1, z_2, \cdots, z_m$ correspond to $w = \infty$; these functions need not be single-valued, but that does not affect our reasoning. The function

* Compare for instance Walsh, *Crelle's Journal*, Vol. 159 (1928), pp. 197-209.

$$r_n(z)/\phi_1(z)\phi_2(z) \cdots \phi_m(z)$$

is analytic in $S(z_1)$ except possibly for branch points, and its modulus is single valued in $S(z_1)$. For z on C, and hence for z on the boundary of $S(z_1)$ inequality (1) is valid, so for z on the boundary of $S(z_1)$ we have

$$(10) \qquad | r_n(z)/\phi_1(z)\phi_2(z) \cdots \phi_m(z) | \leqq M.$$

To be sure, $\phi_i(z)$ is not properly defined on the boundary of $S(z_1)$, but as z approaches that boundary, $| \phi_i(z) |$ approaches unity and $r_n(z)$ is continuous; inequality (10) is to be interpreted in that sense. It follows that (10) is valid throughout the region $S(z_1)$, a region within which the function on the left is analytic except possibly for branch points. Throughout the region common to $S(z_1)$ and C_R we have

$$| \phi_i(z) | \leqq R,$$

so from (10) we derive (2) for z in the region common to $S(z_1)$ and C_R. In a similar manner it is shown that (2) is valid for z in the region common to C_R and *any* $S(z_j)$ containing poles of $r_n(z)$.

Inequality (2) is valid on the boundary of each region $S(z_j)$ containing poles of $r_n(z)$; denote by $\sum$ the set complementary to the totality of such regions $S(z_j)$. Then inequality (2) is valid at each boundary point of $\sum$; the function $\phi_n(z)$ is analytic throughout $\sum$, so inequality (2) is valid at each point of $\sum$, and is valid in particular at each point of C_R belonging to $\sum$. We have already proved the validity of (2) at each point of C_R not in $\sum$, that is at each point of C_R lying in a region $S(z_j)$ containing poles of $r_n(z)$, so the proof of Lemma III is complete.

We add without proof the remarks * that if R' is less than R, the set $C_{R'}$ is interior to the set C_R; if E is an arbitrary closed set interior to C_R, there exists an $R' < R$ such that $C_{R'}$ contains E in its interior; if P represents an arbitrary point interior to C_R, there exists a value of R' less than R such that P lies on the boundary of $C_{R'}$.

Lemma III can now be applied to establish a result analogous to Theorem I; here C is no longer a Jordan region. In the present case *C and D are arbitrary closed point sets with no common point, such that the Dirichlet problem* (for arbitrary continuous boundary values) *has a solution for every region $S(z')$* defined as the totality of points which can be joined to a given point z' of D by broken lines not meeting C.

THEOREM II. *Let $r_n(z)$ be a sequence of rational functions of respective degrees n such that we have for z on C*

$$(8) \qquad | f(z) - r_n(z) | \leqq M/\rho^n, \qquad \rho > 1.$$

* In the proof of these remarks it may be found convenient to use the results of Lebesgue, *Palermo Rendiconti*, Vol. 24 (1907), pp. 371-402.

If the poles of the functions $r_n(z)$ *lie in* D, *then the sequence* $r_n(z)$ *converges uniformly for* z *on* C_R, *where* $R < \rho^{1/2}$.

If (8) *is valid for* z *in* C *and if the function* $r_{n+1}(z) - r_n(z)$ *is of degree* $n+1$, *with its poles in* D, *then the sequence* $r_n(z)$ *converges uniformly for* z *in* C_R, *where* $R < \rho$.

If (8) *is valid for* z *in* C *and if there exists a sequence* $r'_n(z)$ *of functions such that* $r'_n(z) - r_n(z)$ *is a rational function of degree* n *with its poles in* D, *if we have*

$$(9) \qquad | f(z) - r'_n(z) | \leqq M/\rho^n, \qquad z \text{ in } C,$$

and if the sequence $r'_n(z)$ *converges uniformly for* z *in* C_R, *where* $R < \rho$, [R *may or may not depend on the sequence* $r'_n(z)$], *then the sequence* $r_n(z)$ *also converges uniformly for* z *in* C_R.

The proof of Theorem II follows directly that of Theorem I.

III.

If the point set C of Lemma III actually separates the plane, and if the various regions into which C separates the plane are known to contain respectively specific numbers of poles of $r_n(z)$ less than n, then it may occur that a sharper result than Lemma III can be established, with a corresponding application more general than Theorem II. We give only a very simple illustration of this remark:

LEMMA IV. *Let* C *be an arbitrary Jordan curve in whose interior the origin lies. If* $r_{2n}(z)$ *is a rational function of* z *of the form*

$$r_{2n}(z) = a_{-n}z^{-n} + a_{-n+1}z^{-n+1} + \cdots + a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n$$

and if we have for z *on* C

$$(11) \qquad | r_{2n}(z) | \leqq M,$$

then we have also

$$| r_{2n}(z) | \leqq MR^n$$

for z *on* C_R, *where* D *consists of the origin and the point at infinity.*

The proof is essentially contained in the discussion already given in connection with Lemma III.

Lemma IV will be used in proving a result related to Theorem I:

THEOREM III. *Let* C *be an arbitrary Jordan curve in whose interior the origin lies. Let* $r_{2n}(z)$ *be a sequence of rational functions of the form*

$$\begin{aligned} r_{2n}(z) &= r'_n(z) + r''_n(z), \\ r'_n(z) &= a_{0n} + a_{1n}z + a_{2n}z^2 + \cdots + a_{nn}z^n, \\ r''_n(z) &= a_{-1n}z^{-1} + a_{-2n}z^{-2} + \cdots + a_{-nn}z^{-n}, \end{aligned}$$

such that we have for z on C

(12) $$|f(z)-r_{2n}(z)| \leqq M/\rho^n, \qquad \rho>1.$$

Then the sequence $r_{2n}(z)$ converges for z in C_R, uniformly for z in C_{R_1}, $R_1<R$, where $R=\rho$. The point set D is to be taken as consisting of the origin and the point at infinity.

The region C_R is bounded by two curves, $C'_R=C_R(\infty)$ exterior to C and $C_R''=C_R(0)$ interior to C; the sequence $r'_n(z)$ converges for z interior to C'_ρ, uniformly for z interior to C'_σ, $\sigma<\rho=R$, and the sequence $r_n''(z)$ converges for z exterior to C_ρ'', uniformly for z exterior to C_σ'', $\sigma<\rho=R$.

The first part of Theorem III follows directly, for we have by (12) for z on C

$$|f(z)-r_{2n}(z)| \leqq M/\rho^n, \qquad |f(z)-r_{2n+2}(z)| \leqq M/\rho^{n+1},$$
$$|r_{2n+2}(z)-r_{2n}(z)| \leqq (M/\rho^n)(1+1/\rho);$$

and this last inequality yields by Lemma IV, since the function

$$r_{2n+2}(z)-r_{2n}(z)$$

is of the form used in Lemma IV and of degree $2n+2$,

$$|r_{2n+2}(z)-r_{2n}(z)| \leqq (M/\rho^n)(1+1/\rho)R_1^{n+1}$$

for z on C_{R_1}. Thus the sequence $r_{2n}(z)$ converges uniformly for z on C_{R_1} if $R_1<\rho$.

The limit in C_ρ $(\rho=R)$ of the sequence $r_{2n}(z)$ is in C_ρ an analytic function of z, and coincides on C with $f(z)$, so this limit is the analytic extension in C_ρ of $f(z)$ and will be denoted by $f(z)$. The function $f(z)$, defined in the ring bounded by C'_ρ and C_ρ'', may be written in the form

$$f(z)=f_1(z)+f_2(z),$$

where $f_1(z)$ is analytic interior to C'_ρ and $f_2(z)$ is analytic exterior to C_ρ'' and vanishes at infinity: we have

(13) $$f_1(z)=\frac{1}{2\pi i}\int_{C'_\sigma}\frac{f(t)\,dt}{t-z}, \quad \sigma<\rho, \qquad z \text{ interior to } C'_\sigma,$$
$$f_2(z)=\frac{1}{2\pi i}\int_{C_\sigma''}\frac{f(t)\,dt}{t-z}, \qquad z \text{ exterior to } C_\sigma'';$$

these integrals are independent of the precise value of σ. In a similar way we have

(14) $$r'_n(z)=\frac{1}{2\pi i}\int_{C'_\sigma}\frac{r_{2n}(t)\,dt}{t-z}, \qquad z \text{ interior to } C'_\sigma,$$
$$r_n''(z)=\frac{1}{2\pi i}\int_{C_\sigma''}\frac{r_{2n}(t)\,dt}{t-z}, \qquad z \text{ exterior to } C_\sigma''.$$

We have already proved the uniform convergence of $r_{2n}(z)$ to $f(z)$ on C'_σ and C_σ''. Term-by-term integration of the equations just written yields

$$\lim_{n\to\infty} r'_n(z) = f_1(z), \quad \text{uniformly for } z \text{ interior to } C'_\tau, \quad \tau < \sigma < \rho,$$

$$\lim_{n\to\infty} r_n''(z) = f_2(z), \quad \text{uniformly for } z \text{ exterior to } C_\tau''.$$

Theorem III is now completely proved.

Theorem III yields a new result which the reader can readily formulate, by the use of a linear transformation $z = (z' - a)/(z' - b)$.

Theorem III, together with much more general results, has been proved by de la Vallée Poussin * for the case that C is a circle. Approximation on the unit circle by rational functions of the kind indicated is equivalent to approximation by trigonometric sums. Theorem III was previously proved by the present writer (*loc. cit.*) for the case that C is rectifiable; the previous proof is not valid in the present case. We recall also the following complement to Theorem III, which holds whether C is rectifiable or not:

If the function $f(z)$ is analytic in the closed region C_R, then there exist rational functions $r_{2n}(z)$ of the form indicated such that we have (12) *satisfied for $\rho = R$ for z on C.*

We have stated Theorem III for the case that C is a Jordan curve. The reader will notice that Lemma IV, Theorem III, and its complement are still true if C is an arbitrary closed set which separates the plane into two simply connected regions containing respectively the points $z = 0$ and $z = \infty$. Moreover, if we have $\rho > \sigma > 1$, there exists M' such that

$$| f(z) - r_{2n}(z) | \leqq M'\sigma^n/\rho^n, \qquad z \text{ on } C_\sigma.$$

Consequently we have by the use of equations (13) and (14)

$$| f_1(z) - r'_n(z) | \leqq M''\sigma^n/\rho^n, \qquad z \text{ interior to } C'_\tau \qquad \tau < \sigma < \rho,$$
$$| f_2(z) - r_n''(z) | \leqq M''\sigma^n/\rho^n, \qquad z \text{ exterior to } C_\tau''.$$

In particular we have

$$| f_1(z) - r'_n(z) | \leqq M''/\rho_1^n, \qquad z \text{ on } C,$$
$$| f_2(z) - r_n''(z) | \leqq M''/\rho_1^n, \qquad z \text{ on } C,$$

where ρ_1 is an arbitrary number less than ρ and where M'' depends on ρ_1.

IV.

There are various generalizations of our results which can be obtained with little difficulty. We merely mention these and leave the details to the reader.

---

* *Approximation des fonctions* (Paris, 1919), Ch. VIII.

(1) Functions $f(z)$ which are meromorphic instead of analytic on C may be approximated, where some poles of the approximating functions $r_n(z)$ are allowed on C, but with the requirement that the limit points of poles of the functions $r_{n+1}(z) - r_n(z)$ in Theorems I and II and of $r_{2n+2}(z) - r_{2n}(z)$ in Theorem III should lie in D. The results are entirely analogous to those already obtained.

(2) The properties of a sequence $r_n(z)$ and its limit often depend not primarily on the position of the poles of $r_n(z)$ but on the limiting position of the poles of $r_{n+1}(z) - r_n(z)$, that is on the limit points of those poles. A correspondingly more general statement of our results can be formulated.

(3) As has already been suggested, Lemma IV and Theorem III are merely special cases of more general results, where C separates the plane or not and D consists of several separated parts. Results exist more general than Lemma III and Theorem II, if the various parts of D are known to contain respectively a specific number (less than n) of poles of $r_n(z)$ or of $r_{n+1}(z) - r_n(z)$. This is true, moreover, whether or not parts of D are separated by C.

(4) Throughout the present paper our measure of the approximation of $r_n(z)$ to $f(z)$ has been

$$\max | f(z) - r_n(z) |, \qquad z \text{ on } C.$$

There exist other interesting measures of the approximation of $r_n(z)$ to $f(z)$ on C, such as (i) the line integral

$$\int | f(z) - r_n(z) |^p \quad | dz |, \qquad p > 0,$$

extended over the boundary of C, if that boundary is rectifiable; (ii) the surface integral

$$\int\int | f(z) - r_n(z) |^p \quad dS, \qquad p > 0,$$

extended over the area of C; (iii) the integral

$$\int | f(z) - r_n(z) |^p \quad | dw |, \qquad p > 0,$$

extended over the circle $\gamma : | w | = 1$ after mapping of C or of such a region as $S(z')$ onto the interior or exterior of γ. Any of these new measures of approximation may be used, even with the introduction of a positive weight or norm function, and under suitable restrictions there can be proved results analogous to those we have here established in detail.

(5) The approximation of harmonic functions by harmonic rational functions may be studied by our present methods.

ON APPROXIMATION TO A MAPPING FUNCTION BY POLYNOMIALS.

By O. J. FARRELL.*

1. *Introduction.* In a paper on approximation to an arbitrary function of a complex variable by polynomials, Professor Walsh † raised, among others, the question: *What conformal maps can be approximately performed by means of polynomials with an error uniformly and arbitrarily small?* The present paper contributes a new result in connection with this question.

We confine attention to the conformal mapping of a finite simply connected region on the interior of the unit circle. For such a map the question raised is answered by the following theorem:

THEOREM I. *Let G be a finite simply-connected region in the z-plane, and let $f(z)$ be a function which maps G conformally on the interior of the unit circle. Then in order that $f(z)$ when suitably defined on the boundary γ of G be continuous on $G+\gamma$ and be also uniformly developable on $G+\gamma$ in a series of polynomials in z, it is necessary and sufficient that every point of γ be a simple* ‡ *boundary point of G and that γ be also the boundary of an infinite region.*

Theorem I is the chief result of the present paper; most of the paper is devoted to its proof. One further result, obtained by a simple application of Theorem I, may be mentioned here:

THEOREM II. *In the z-plane let G be a finite simply connected region whose boundary γ consists wholly of simple boundary points of G and is also the boundary, not of an infinite region, but of a finite region R distinct from G. Let $f(z)$ be a function which maps G conformally on the interior of the unit circle. Then $f(z)$ may be uniformly expanded on $G+\gamma$ in a series of polynomials in $1/(z-\alpha)$, where $z=\alpha$ is any chosen fixed point in R.*

An example of a region to which Theorem II is applicable is the region

---

* National Research Fellow.

† *Transactions of the American Mathematical Society*, Vol. 30 (1928), pp. 472-482.

‡ That is, contained in just one boundary element (Primende). See Carathéodory, "Über die Begrenzung einfach zusammenhängender Gebiete," *Mathematische Annalen*, Vol. 73 (1913), pp. 323-370, § 44.

consisting of a strip closed at one end which lies interior to the unit circle and which winds around infinitely often, approaching every point of the circle as it does so.

It is interesting to note in this connection the existence of a region G whose boundary γ although made up wholly of simple boundary points is not the boundary of any other region either finite or infinite. Consider, for instance, the region constructed by Professor Carathéodory* to show that boundary elements of the second kind can occur with the power of the continuum in every interval of boundary elements. Let P be any boundary point of this region that lies in the interior of the rectangle $ABCD$. Then it is readily seen that in a sufficiently small neighborhood of P are to be found only points of $G+\gamma$. This means of course that the complete boundary γ of G does not serve as the boundary of any region other than G.

2. *Proof of necessity of conditions of Theorem I.* The fact that the boundary γ of G must consist entirely of simple boundary points of G results from the following

LEMMA. *Let $w=f(z)$ be a function which maps a finite simply connected region G of the z-plane conformally on the region $|w|<1$. A necessary and sufficient condition that $f(z)$ be definable on the boundary γ of G so as to be continuous on $G+\gamma$ is that every point of γ be a simple boundary point of G.*

The proof of this lemma is immediate by Theorems XIII and XXI of the paper by Carathéodory already referred to above.

It remains to show that γ must also be the boundary of an infinite region. Suppose this were not so. All points of the plane which can be joined with the point $z=\infty$ by a Jordan arc containing no point of γ form a region whose boundary β consists wholly of points of γ, but which by hypothesis does not exhaust the point set γ. Let z_0 be any fixed point of G. All points which can be joined to z_0 by a Jordan arc containing no points of β form a region T bounded by β and containing as interior points those points of γ which do not belong to β.

If now there exists a sequence of polynomials converging uniformly to $f(z)$ on $G+\gamma$, the sequence will in particular converge uniformly on β and hence also on $T+\beta$. This sequence thus defines a function regular-analytic in T, continuous on $T+\beta$, and assuming its maximum absolute value, namely unity, at interior points of T. But this is impossible. Hence

* *Loc. cit.*, § 49.

the complete boundary γ of G must also constitute the boundary of an infinite region.

3. *A theorem on the convergence of a sequence of mapping functions.* We require for the proof of the sufficiency of the conditions of Theorem I the following theorem:

THEOREM III. *In the z-plane let G be a finite simply connected region containing the origin whose boundary γ consists wholly of simple boundary points of G and is also the boundary of an infinite region. Let*

$$G_1, G_2, G_3, \cdots G_n, \cdots \tag{1}$$

denote a sequence of uniformly bounded simply connected regions such that the closed region $G+\gamma$ lies interior to every G_n. Let $w = f_1(z)$, $f_2(z)$, $f_3(z), \cdots$; $f(z)$ denote the functions that map conformally $G_1, G_2, G_3, \cdots$; G respectively on the region $|w| < 1$ so that in every case the points $z = 0$ and $w = 0$ together with the directions of the positive axes of reals at these points correspond to each other. Suppose that $f(z)$ has been defined on γ so as to be continuous on $G+\gamma$. Then if G be the kernel of the sequence (1) and if this sequence converge to its kernel, we have*

$$\lim_{n\to\infty} f_n(z) = f(z) \tag{2}$$

uniformly for all z on $G+\gamma$.

It follows from the hypotheses of Theorem III by results of Bieberbach † that (2) holds uniformly in any closed region lying wholly interior to G. We shall show, however, that (2) holds uniformly for all z in the open region G. From this it will then follow, due to the continuity of $f(z)$ and $f_n(z)$ on $G+\gamma$, that (2) holds uniformly for all z on $G+\gamma$.

Suppose that (2) does not hold uniformly for all z in the open region G. Then there exists a sequence of points of G

$$z_1, z_2, \cdots z_k, \cdots \tag{3}$$

together with a subsequence of functions

$$f_{\mu_1}(z), f_{\mu_2}(z), \cdots f_{\mu_k}(z), \cdots \tag{4}$$

and a positive number Δ such that

$$|f_{\mu_k}(z_k) - f(z_k)| > \Delta, \qquad (k = 1, 2, 3, \cdots). \tag{5}$$

* See Bieberbach, *Lehrbuch der Funktionentheorie*, Vol. 2 (1927), pp. 12-13.

† *Loc. cit.*, p. 13. Compare Walsh, *loc. cit.*, p. 474.

No interior point of G can be a limit point of (3) because of the uniform convergence of $f_n(z)$ in every closed region that lies in G. We may assume, then, that (3) converges in the sense of Carathéodory * to a boundary element E_0 of G, since in any event a subsequence can be chosen which has this property. Let w_0 designate the point on the circle $|w| = 1$ which corresponds by $w = f(z)$ to E_0.

The boundary element E_0 may be defined by a chain of cross-cuts $q_1, q_2, \cdots$ which lie on concentric circles and converge to a point of E_0. Let $g_1, g_2, \cdots$ denote the corresponding chain of subregions.

A point in G_n which can be connected to a point on the boundary γ_n of G_n by a Jordan arc J of diameter less than δ is transformed by $w = f_n(z)$ into a point whose distance from the point corresponding by this same relation to the boundary element of G_n defined by J is less than a quantity $\theta(\delta)$ which is independent of n and approaches zero with δ.†

Let an arbitrary positive ϵ be given. Choose δ so that $\theta(\delta) < \epsilon$. From the cross-cuts $q_1, q_2, \cdots$ choose one q_N whose length is less that $\delta/3$, whose transform by $w = f(z)$ lies entirely within distance ϵ of w_0, and such that the point $z = 0$ is exterior to g_N. Let z' designate the middle point of q_N. Denote by Q_1 and Q_2 the end-points ‡ of q_N. Choose a positive number η less than $\delta/3$ and also less than $\overline{Q_1Q_2}/2$.

By choosing n_0 sufficiently large, we have for all $n \geqq n_0$

(a) $|f_n(z') - f(z')| < \epsilon$,

(b) within distance η of every point of γ lie points of γ_n.

The condition (a) can be obtained by virtue of the result of Bieberbach mentioned above. The fact that (b) can also be obtained may be seen as follows. If this were not possible, there would exist a sequence of points $P_1, P_2, P_3, \cdots$ on γ together with a subsequence of regions $G_{\mu_1}, G_{\mu_2}, G_{\mu_3}, \cdots$ such that for every k no point of the boundary γ_{μ_k} of G_{μ_k} lies within distance η of P_k. We may assume that the sequence $P_1, P_2, \cdots$ has a unique limit point P_0. Then all points P_k for sufficiently large k lie within distance $\eta/2$ of P_0. Hence no boundary γ_{μ_k} from a certain k on comes within distance $\eta/2$ of P_0.

* *Loc. cit.*, definition No. III, top of p. 332; also Theorem VII, p. 341.

† See Lindelöf, "Sur un principe général de l'analyse," *Acta Societatis Scientiarum Fennicae*, Vol. 46, No. 4, the two italicized results on page 16. Compare Walsh, *loc. cit.*, p. 475.

‡ The points Q_1 and Q_2 are distinct. If this were not so, points of γ and hence also points of the infinite region F bounded by γ could be found both inside and outside the circle formed by q_N. But this would be absurd, for F would not then be a region.

But then G would not be the kernel of the subsequence $G_{\mu_1}, G_{\mu_2}, \cdots$, contrary to the hypothesis that the sequence (1) converges to its kernel G.

Consider any region G_n of (1) for which $n \geqq n_0$. Join Q_1 and Q_2 each with the nearest point of γ_n by means of the straight line segments $\overline{Q_1T_1}$ and $\overline{Q_2T_2}$. Let R_1 and R_2 denote the first points of intersection of $\overline{Q_1T_1}$ and $\overline{Q_2T_2}$ respectively with q_N when one proceeds from T_1 to Q_1 and from T_2 to Q_2. (It can happen of course in simple cases that R_1 coincides with Q_1, or R_2 with Q_2, or both.) Thus we obtain a cross-cut $T_1R_1R_2T_2$ of G_n consisting of the straight line segments $\overline{T_1R_1}$ and $\overline{R_2T_2}$ together with that portion of q_N which joins R_1 and R_2. Call this cross-cut σ. We notice that σ contains z' and that the length of σ is less than δ.

Denote by G'_n and G_n'' the two sub-regions of G_n defined by σ. For sufficiently small choice of η the subregion g_{N+1} of G is wholly interior to one of the two regions G'_n and G_n'', say G'_n, while the point $z = 0$ is interior to the other region G_n''. This may be seen as follows. About z' as center draw a circle K lying entirely interior to G. The cross-cut q_N divides the interior of K into two regions H' and H'' such that one of them, say H', is entirely interior to g_N and H'' is wholly exterior to g_N. Join an arbitrary fixed point of H' with an arbitrary fixed point of g_{N+1} by a Jordan arc lying in g_N. Join also a fixed point of H'' with the point $z = 0$ by a Jordan arc lying in G but exterior to g_N. Let μ denote the minimum distance between a point of either arc and the boundary γ of G. Let ρ denote the minimum distance between either of the points Q_1, Q_2 and a point of K. Finally let λ denote the minimum distance * between q_N and the closed region consisting of g_{N+1} and its boundary. Then it will be seen that we need merely to choose any positive η which (in addition to being less than $\delta/3$ and $\overline{Q_1Q_2}/2$ as already required above) is less than the least of the positive numbers μ, ρ, λ.

The transform τ of σ by $w = f_n(z)$ is a cross-cut of the region $|w| < 1$ lying entirely within distance $\theta(\delta) < \epsilon$ of the point on the circle $|w| = 1$ which corresponds by $w = f_n(z)$ to the boundary element of G_n defined by the segment $\overline{Q_1T_1}$. Thus every point of τ is within distance 2ϵ of the point $w = f_n(z')$. It follows then from the inequalities

$$|f_n(z') - f(z')| < \epsilon \quad \text{and} \quad |f(z') - w_0| < \epsilon$$

that all points of τ are within distance 4ϵ of w_0. Hence one of the two subregions of $|w| < 1$ defined by τ must lie wholly within distance 4ϵ of w_0.

* This minimum distance λ is positive. In the contrary case, an end-point of q_N, say Q_1, would have to lie on the boundary of g_{N+1}. But then Q_1 would not be a simple boundary point of G. See Carathéodory, *loc. cit.*, Theorem XXII.

And this region is the transform by $w = f_n(z)$ of G'_n, at least for $\epsilon < \frac{1}{4}$, since then the point $w = 0$, which corresponds to $z = 0$, is a point of the transform of the other region G_n''.

The transform of g_{N+1} by $w = f_n(z)$ is therefore a region every point of which is within distance 4ϵ of w_0. And the transform of g_{N+1} by $w = f(z)$ lies wholly within distance ϵ of w_0, since g_{N+1} is a part of g_N. And so we have for $\epsilon < 5\Delta$

$$| f_n(z) - f(z) | < \Delta$$

for all z in g_{N+1} and for all $n \geqq n_0$. But for sufficiently large k the points z_k of (3) are interior points of g_{N+1}. And for these points (5) holds. Here we have a contradiction, and hence it must be true that (2) holds uniformly in the open region G and consequently also on $G + \gamma$.

4. *Proof of sufficiency of conditions of Theorem I.* It is no loss of generality to assume that the given region G contains the origin and that the mapping makes correspond the points $z = 0$ and $w = f(z) = 0$ together with the directions of the positive axes of reals at these points; for this may always be effected by a preliminary linear transformation of the form

$$\zeta = az + b. \tag{6}$$

Then $f(z)$ becomes $g(\zeta)$, and an expansion of $g(\zeta)$ in the transform of $G + \gamma$ by (6) in terms of polynomials in ζ is equivalent to an expansion of $f(z)$ on $G + \gamma$ in terms of polynomials in $az + b$, that is, in terms of polynomials in z.

Choose a sequence of uniformly bounded simply connected regions $G_1, G_2, \cdots$ which have the given region G as their kernel and which converge to their kernel. Let $w = f_1(z), f_2(z), \cdots$ denote the functions which map these regions respectively on the interior of the unit circle so that in every instance the points $z = 0$ and $w = 0$ correspond, together with the directions of the positive axes of reals at these points.

Let an arbitrary positive ϵ be given. Choose by virtue of Theorem III a function $f_n(z)$ such that

$$| f(z) - f_n(z) | < \epsilon/2, \qquad z \text{ on } G + \gamma. \tag{7}$$

By Runge's well known theorem * there exists a polynomial $P(z)$ such that

$$| f_n(z) - P(z) | < \epsilon/2, \qquad z \text{ on } G + \gamma. \tag{8}$$

Combining (7) and (8) we have the desired result

* See, for instance, Bieberbach, *loc. cit.*, Vol. 1 (1921), p. 296.

$$|f(z) - P(z)| < \epsilon, \qquad z \text{ on } G + \gamma.$$

5. *Two applications of Theorem I.* Our first application of Theorem I is Theorem II, the proof of which is very simple. Let $z = \alpha$ be an arbitrarily chosen fixed point in R, and let the circle C: $|z - \alpha| = r$ lie wholly in R. The transformation

$$\zeta = r^2/(z - \alpha) + \alpha \tag{9}$$

exchanges the interior of C with its exterior. Let $g(\zeta)$ denote the inverse of (9). Then by Theorem III the function $f[g(\zeta)]$, which maps the transform of G by (9) on the region $|w| < 1$, may be uniformly approximated as closely as desired in this closed region by a polynomial in ζ. But this means that $f(z)$ can be uniformly approximated as closely as desired on $G + \gamma$ by a polynomial in $1/(z - \alpha)$.

Another simple application of Theorem I results in

THEOREM IV. *In the z-plane let G be a region whose boundary γ consists wholly of simple boundary points of G and is also the boundary of an infinite region. If the function $F(z)$ is analytic in G, continuous on $G + \gamma$, and constant on each boundary element of G, then $F(z)$ may be uniformly approximated as closely as desired on $G + \gamma$ by a polynomial in z.*

For convenience in proving this theorem we require the following

LEMMA. *The function $f^n(z)$, where n is any positive integer and $f(z)$ is the mapping function of Theorem I, may be uniformly approximated as closely as desired on $G + \gamma$ by a polynomial in z.*

Let an arbitrary positive ϵ be assigned. Choose by virtue of Theorem I a polynomial $P(z)$ such that

$$|P(z) - f(z)| < \frac{\epsilon}{n(1+\epsilon)^{n-1}}, \qquad z \text{ on } G + \gamma.$$

Then we have

$$\begin{aligned} |P^n(z) - f^n(z)| &= |P(z) - f(z)| \\ &\quad \cdot |P^{n-1}(z) + P^{n-2}(z)f(z) + P^{n-3}(z)f^2(z) + \cdots + f^{n-1}(z)| \\ &< \frac{\epsilon}{n(1+\epsilon)^{n-1}} \cdot n(1+\epsilon)^{n-1} = \epsilon. \end{aligned}$$

Let $z = g(w)$ be the inverse of $w = f(z)$, where $f(z)$ is a function which maps the given region G of Theorem IV on the region $|w| < 1$. Then for $|w| < 1$ we have

$$F(z) = F[g(w)] = H(w),$$

and $H(w)$ is analytic for $|w| < 1$. Moreover, $H(w)$ takes on a boundary value at each point of the circle $|w| = 1$. Let $w = w_0$, for instance, be a point of this circle, and let E_0 denote the boundary element of G that corresponds by $z = g(w)$ to w_0. If $w_1, w_2, \cdots$ be any sequence of points interior to $|w| < 1$ converging to w_0, the corresponding sequence $z_1, z_2, \cdots$ $(z_n = g(w_n))$ converges to E_0. But then the sequence

$$H(w_1), H(w_2), \cdots,$$

or what is the same,

$$F(z_1), F(z_2), \cdots,$$

converges to a definite value, namely the constant value of $F(z)$ on E_0. So if we define $H(w)$ on $|w| = 1$ as equal to its boundary values there, $H(w)$ is continuous for $|w| \leqq 1$ as well as analytic for $|w| < 1$.

Then, given an arbitrary positive ϵ, we can find a polynomial * $P(w)$ such that

$$|H(w) - P(w)| < \epsilon/2, \qquad |w| \leqq 1.$$

That is

$$|F(z) - P[f(z)]| < \epsilon/2, \qquad z \text{ on } G + \gamma.$$

But by the lemma each positive integral power of $f(z)$ can be approximated as closely as desired on $G + \gamma$ by a polynomial in z. Consequently a polynomial $p(z)$ exists such that

$$|P[f(z)] - p(z)| < \epsilon/2, \qquad z \text{ on } G + \gamma.$$

Combining the last two inequalities we have

$$|F(z) - p(z)| < \epsilon, \qquad z \text{ on } G + \gamma.$$

University of Munich,
Munich, Germany.

* See Walsh, *loc. cit.*, Theorem I.

ON INFINITE ORTHOGONAL MATRICES.

By Monroe Harnish Martin.

TABLE OF CONTENTS.

INTRODUCTION.

If one interprets the elements m_{jk} $(j, k = 1, 2, \cdots n < +\infty)$ of an arbitrary real or complex n rowed square matrix $\boldsymbol{M} = \| m_{jk} \|$ as the coördinates of a point in a Euclidean space of n^2 real or complex dimensions, the set of points representing the real unitary matrices (i. e. the orthogonal matrices) divides itself into two separate continua, one of which contains the rotation matrices (i. e. the orthogonal matrices of determinant $+1$) and the other the reflection matrices (i. e. the orthogonal matrices of determinant -1). The interpretation of a finite matrix as a point in a Euclidean space of a suitable number of dimensions may obviously be extended to infinite matrices by employing an abstract manifold of ∞^2 real or complex dimensions. The aim of the present paper is to obtain a division of this abstract manifold into two sub-manifolds in a manner which is a natural extension of the afore-mentioned division of the space of finite orthogonal matrices. The procedure is to first obtain a one-to-one parametric representation of the entire manifold of the infinite orthogonal matrices which divides the ortho-

gonal matrices into four distinct classes, and then to show how the matrices of these four classes may be reapportioned into two sub-manifolds of the manifold of the infinite orthogonal matrices, the two sub-manifolds coinciding, for finite orthogonal matrices, with the two separate continua noted above. Since the formula of Cayley does not yield, as we show in some detail in the next few pages, a parametric representation of the entire manifold of the orthogonal matrices, it has been found necessary to employ the spectral theory which yields a one-to-one exponential representation of the entire manifold of the orthogonal matrices. Throughout this paper we employ bold-faced capital letters to designate finite matrices, the Roman capitals being retained for infinite matrices.

We denote by

$$\Phi[\mathrm{M}; x] = \sum_{p=1}^{+\infty} \sum_{q=1}^{+\infty} m_{pq} x_p \bar{x}_q, \tag{1}$$

the *compound form* (Kopplungsform) of an arbitrary bounded † matrix $\mathrm{M} = \| m_{pq} \|$ and shall call a real or complex vector $x = (x_1, x_2, \cdots)$ a *unit vector* if

$$| x | = (\Phi[\mathrm{E}; x])^{1/2} = \left(\sum_{p=1}^{+\infty} | x_p |^2\right)^{1/2} = 1, \tag{2}$$

in which E is the unit matrix. The compound forms of bounded matrices possess the obvious uniqueness property in that the necessary and sufficient condition for the equality

$$\Phi[\mathrm{M}; x] = \Phi[\mathrm{N}; x]$$

to hold for every unit vector x is that $\mathrm{M} = \mathrm{N}$. Expressed otherwise, the equality

$$\Phi[\mathrm{M}; x] = 0,$$

holds for every unit vector x then and only then if M is the zero matrix $\| 0 \|$.

We shall understand by the *domain* ‡ $\Delta(\mathrm{M})$ of an arbitrary bounded matrix M the point-set which one constructs in the complex plane from the values, together with their limiting values, which are taken by the compound form (1) for those vectors x satisfying (2), i. e. for the totality of the unit vectors which compose the so-called complex Hilbert sphere (2). The domain $\Delta(\mathrm{M})$ of a bounded matrix M is then, from its definition, a bounded closed point set in the complex plane.

---

† D. Hilbert, "Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen," (1912), p. IV. This book will be referred to as "Hi".

‡ Termed by German writers "Wertevorrat."

According to a theorem of Toeplitz † there exists for every bounded matrix M at most one bounded matrix K for which the two independent conditions $KM = E$, $MK = E$ are fulfilled. If such a bounded matrix exists we designate it as the reciprocal of M and denote it by M^{-1}. We then understand by the *spectrum* ‡ of a bounded matrix N the set of those points λ in the complex plane for which $(\lambda E - N)^{-1}$ does not exist and denote this point-set by $\Lambda(N)$. It is known § that the point-set $\Lambda(N)$ thus defined for every bounded matrix N contains at least one point and is a closed subset of the domain $\Delta(N)$. If we denote by $\boldsymbol{l}[M]$ the greatest lower bound of the compound form $\Phi[N; x]$ of a bounded matrix N under the condition that $|x| = 1$ and by N^* the transposed matrix of the conjugate complex elements of N (i. e. $N^* = \bar{N}'$), the number λ is then and only then, according to Toeplitz,¶ in the spectrum of N if at least one of the following two equations is fulfilled

$$(3) \quad \boldsymbol{l}[(\lambda E - N)(\lambda E - N)^*] = 0, \qquad \boldsymbol{l}[(\lambda E - N)^*(\lambda E - N)] = 0.$$

If N is commutable with N^*, i. e. N is a normal matrix, the two conditions in (3) are identical.

The subset of the spectrum composed of those and only those points λ for which the homogeneous equations belonging to the bounded matrix $\lambda E - N$ possess a solution x for which $|x| = 1$ is called the *point spectrum* ‖ of N. The points of the spectrum do not necessarily all belong to the point spectrum and it is possible that the point spectrum of N contains no single point. It is clear that the spectrum of $cE - N$ is obtained from the spectrum of N by a translation, determined by the real or complex number c, in the complex plane. If T is a bounded matrix for which T^{-1} exists then the bounded matrices M and TMT^{-1} possess the same spectrum and the same point spectrum.

A matrix N is said to be *Hermitian* if $N^* = N$, *unitary* if N^{-1} exists and is equal to N^*, and *skew-symmetric* if $N' = -N$. It is known that the unitary matrices form a group of bounded matrices and that no unitary

† O. Toeplitz, "Die Jacobische Transformation der quadratischen Formen von unendlichvielen Veränderlichen," *Göttingen Nachrichten* (1907), pp. 101-109. This paper will be referred to as "T".

‡ A. Wintner, "Zur Theorie der beschränkten Bilinearformen," *Mathematische Zeitschrift*, Vol. 30 (1929), p. 239. This paper will be referred to as "W".

§ See "W", pp. 241-243.

¶ See "T".

‖ This definition of the point spectrum of an arbitrary bounded matrix N coincides for bounded Hermitian and unitary matrices with the usual definition for the point spectrum of matrices of these two classes.

matrix is completely continuous (vollstetig). The domain of an arbitrary bounded matrix N is a unitary invariant, i.e. if U is any unitary matrix, then $\Delta(\mathrm{N}) = \Delta(\mathrm{UNU}^{-1})$.

The domain $\Delta(\mathrm{H})$ of a bounded Hermitian matrix H is a bounded, closed interval of the real axis of the complex plane, the end-points of which always belong † to the spectrum $\Lambda(\mathrm{H})$ of H. Since the spectrum of any bounded matrix, as has been mentioned above, is a subset of the domain of the matrix, the spectrum $\Lambda(\mathrm{H})$ of a Hermitian matrix H lies entirely on the real axis of the complex plane. The domain of a unitary matrix U is a convex subset of the unit circle, the spectrum of U being contained on the boundary of the unit circle. The real subset of the manifold of the unitary matrices forms the manifold of the *orthogonal matrices* O so that the spectrum of O is likewise contained on the boundary of the unit circle and is in addition symmetric with respect to the axis of reals inasmuch as the spectrum of any real bounded matrix possesses this symmetry property. The domain of a bounded real skew-symmetric matrix is contained on the axis of imaginaries and, as may easily be verified directly, is symmetric with respect to the axis of reals. For our purpose it will be more instructive to derive this symmetry property of the domain of a bounded real skew-symmetric matrix as follows: The matrix $i\mathrm{S}$ in which S is a bounded real skew-symmetric matrix is obviously Hermitian so that its domain and, therefore, its spectrum is contained on the axis of reals. Since the spectrum of S is symmetric with respect to the axis of reals, the spectrum of $i\mathrm{S}$ is symmetric with respect to the axis of imaginaries and, therefore, also the domain of $i\mathrm{S}$ is symmetric with respect to the axis of imaginaries inasmuch as, as has been mentioned above, the end-points of the domain of a bounded Hermitian matrix H always belong to the spectrum of H. Finally, since $\Phi[\mathrm{S}; x] = -i\Phi[i\mathrm{S}; x]$, it follows that the domain of S is symmetric with respect to the axis of reals.

A pair of matrices (M, N) is said to be *separated* if the elements m_{pq} of M and the elements n_{pq} of N stand in the following relationship to one another:

$$(4) \qquad \begin{aligned} &\text{if} \quad m_{pq} \neq 0 \quad \text{then} \quad n_{pk} = n_{jp} = n_{jq} = n_{qk} = 0, \\ &\text{if} \quad n_{pq} \neq 0 \quad \text{then} \quad m_{pk} = m_{jp} = m_{jq} = m_{qk} = 0, \end{aligned}$$

in which the two subscripts j, k take independently (p and q fixed) all positive integral values.

† F. Riesz, "Les systèmes d'équations linéaires a une infinité d'inconnues" (1913), pp. 139-140. For a direct demonstration cf. A. Wintner, "Spektraltheorie der unendlichen Matrizen" (1929), pp. 146-148. The book of Riesz will be referred to as "R" and that of Wintner as "S".

If two bounded matrices M and N form a separated pair (M, N) and if Φ_M Φ_N are any two points whatsoever in the domains $\Delta(M)$ and $\Delta(N)$ respectively, it is clear from (4) that there exist two sequences of unit vectors so that (cf. the definition of the domain of a bounded matrix on p. 580)

$$(5)\quad \begin{aligned} \Phi_M &= \lim_{k\to+\infty} \Phi[M; y_k] \quad \text{while} \quad 0 = \Phi[N; y_k] \quad \text{for} \quad k = 1, 2, \cdots, \\ \Phi_N &= \lim_{k\to+\infty} \Phi[N; z_k] \quad \text{while} \quad 0 = \Phi[M; z_k] \quad \text{for} \quad k = 1, 2, \cdots, \end{aligned}$$

It is not difficult to see that if the matrix pair (M, N) is separated, the infinitely many matrix pairs

$$(6)\qquad (M^p, N^q); \qquad (p, q = 1, 2, \cdots)$$

are also separated. A pair of bounded matrices (M, N) is said to be *perpendicular* if

$$(7)\qquad MN = NM = \| 0 \|,$$

and it is easy to see that while a separated pair of bounded matrices is always perpendicular, it does not follow that a perpendicular pair of matrices is necessarily separated. In addition the property of a pair of bounded matrices to be perpendicular is readily seen, in contrast to the property of a matrix pair to be separated, to be invariant with respect to simultaneous matrix transformations of the two matrices composing the pair. It is trivial from (7) that for a perpendicular matrix pair (M, N) we have

$$(8)\qquad (M + N)^n = M^n + N^n; \qquad (n = 1, 2, \cdots).$$

We employ (5), (6) and (8) in order to show that a separated pair of Hermitian matrices (M, N) possesses in the notation of p. 581 the property

$$(9)\qquad \Lambda(M + N) = \Lambda(M) + \Lambda(N)$$

in which the $+$ sign denotes the usual process of the addition of two point sets, i. e. for example $\Lambda(M) + \Lambda(M) = \Lambda(M)$. In order to prove (9) let μ and ν be any two points (not necessarily different) in the respective spectra of the bounded Hermitian matrices M and N forming the separated matrix pair (M, N). According to (3) there accordingly exist two sequences $\{y_n\}$ and $\{z_n\}$ of unit vectors for which

$$(10)\quad \begin{aligned} &\lim_{n\to+\infty} \Phi[(\mu E - M)(\mu E - M)^*; y_n] \\ &\qquad\qquad = \lim_{n\to+\infty} \Phi[(\mu^2 E - 2\mu M + M^2); y_n] = 0, \\ &\lim_{n\to+\infty} \Phi[(\nu E - N)(\nu E - N)^*; z_n] \\ &\qquad\qquad = \lim_{n\to+\infty} \Phi[(\nu^2 E - 2\nu N + N^2); z_n] = 0, \end{aligned}$$

and we now employ (5), (6) and (8) to show that both μ and ν lie in the spectrum of the Hermitian matrix $M + N$. In order that a point λ lies in the spectrum of $M + N$ it is, again according to (3), necessary and sufficient that there exist a sequence of unit vectors $\{x_n\}$ for which

$$(11) \quad \lim_{n\to+\infty} \Phi\,[(\lambda E - M - N)(\lambda E - M - N)^*; x_n] \\ = \lim_{n\to+\infty} \Phi\,[(\lambda^2 E - 2\lambda M - 2\lambda N + M^2 + N^2); x_n] = 0,$$

inasmuch as from (8) we have $(M + N)^2 = M^2 + N^2$. Since, according to (6), the matrix pairs (M, N), (M, N^2), (M^2, N), (M^2, N^2) are all separated, the condition (11) may obviously, from (10), be satisfied by putting either

$$\lambda = \mu \text{ and } \{x_n\} = \{y_n\} \quad \text{or} \quad \lambda = \nu \text{ and } \{x_n\} = \{z_n\},$$

inasmuch as, as follows from (5) and (6), the sequences $\{y_n\}$ and $\{z_n\}$ in (10) may be chosen to have

$$\Phi\,[M; z_n] = 0,\ \Phi\,[M^2; z_n] = 0; \quad \Phi\,[N; y_n] = 0,\ \Phi\,[N^2; y_n] = 0,$$

for $n = 1, 2, \cdots$, i. e. the points μ and ν both belong to the spectrum of $M + N$.

An important application for p. 607 of the text concerns itself with the special case in which M is a diagonal matrix, the elements of which are either 0 or $+1$. The complementary matrix $E - M$ is accordingly of the same type as M and forms with M the separated pair $(E - M, M)$. We then construct the separated pair $(c\,[E - M] + N, cM)$ where c is any number, real or complex, and apply the result (9) of the preceding paragraph to this matrix pair, thereby obtaining

$$\Lambda(cE + N) = \Lambda(c\,[E - M] + N + cM) = \Lambda(c[E - M] + N) + \Lambda(cM),$$

or, in particular for $c = \pi$ and $N = iS$

$$(12) \qquad \Lambda(\pi M) + \Lambda(\pi\,[E - M] + iS) = \Lambda(\pi E + iS).$$

Since the spectrum $\Lambda(\pi M)$ of πM only contains the two points 0 and π, it is clear from (12) that $\lambda > \pi$ is then and only then a point in the spectrum of $\pi\,[E - M] + iS$ if it is a point in the spectrum of $\pi E + iS$. It is furthermore trivial that 2π is then and only then in the point spectrum of $\pi E + iS$ if $+\pi$ (and if S is real and skew-symmetric then also $-\pi$) is a point in the point spectrum of iS so that we are led to the following result employed on p. 610 of the text: If S is a real skew-symmetric matrix whose spectrum is contained in the closed interval $[-i\pi, +i\pi]$ and if M is a diagonal matrix, the elements of which are either 0 or $+1$, forming with S the

separated matrix pair (M, S), then the point 2π lies in the point spectrum of the Hermitian matrix $\pi(E-M)+iS$ if, and only if, if the end-points $-\pi$ and $+\pi$ of the interval $[-\pi, +\pi]$ belong to the point spectrum of iS.

A matrix I is said to be an *idempotent* matrix if $I^2 = I$ (inasmuch as then $I^n = I$ for $n = 1, 2, \cdots$). The Hermitian idempotent matrices are designated by Hilbert as "Einzelmatrizen." The bounded idempotent matrices I possess the property that the continuously many matrices $2\pi i n I$ ($n = 0, \pm 1, \cdots$) are all "logarithms" of E, i. e. $\exp(2\pi i n I) = E$ [Cf. (80)]. Since the matrix

$$[(\lambda - 1)E + I]/[\lambda(\lambda - 1)]$$

represents, as may easily be verified, for $\lambda(\lambda-1) \neq 0$ the reciprocal matrix of $(\lambda E - I)$ for every bounded idempotent matrix I, the spectrum of I contains at most the two points $\lambda = 0$ and $\lambda = +1$. For real symmetric (and therefore bounded) idempotent matrices there exists, according to Hilbert,† an orthogonal matrix O so that $OIO^{-1} = \| \kappa_{pq} \|$ where $\kappa_{pq} = 0$ for $p \neq q$ and $\kappa_{pp} = 0$ or $\kappa_{pp} = +1$ (Cf. the matrices M and $E - M$ above). We denote by α' the number of times 0 occurs in the sequence $\{\kappa_{pp}\}$ and by β' the number of times $+1$ occurs in this sequence so that $\alpha' \geqq 0$, $\beta' \geqq 0$ and $\alpha' + \beta' = +\infty$. The numbers α' and β' are uniquely determined by the matrix I, i. e. they are independent of the special choice of the matrix O employed in transforming I into the diagonal form $\| \kappa_{pq} \|$. It will be necessary to distinguish the following four classes of idempotent matrices:

$$\begin{array}{rlll} \text{I}: & \alpha' = +\infty, & \beta' = 2m & \text{for } 0 \leqq m < +\infty, \\ \text{II}: & \alpha' = +\infty, & \beta' = 2m+1 & \text{for } 0 \leqq m < +\infty, \\ \text{III}: & \alpha' = m, & \beta' = +\infty & \text{for } 0 \leqq m < +\infty, \\ \text{IV}: & \alpha' = +\infty, & \beta' = +\infty. & \end{array}$$

We call a matrix Q a *square root of E* if $Q^2 = E$. The bounded square roots of E may be characterized by the matrix equation $Q^{-1} = Q$ and possess the property that the continuously many matrices $2\pi i n Q$ ($n = 0, \pm 1, \pm 2, \cdots$) are all "logarithms" of E (Cf. above). The general real symmetric square root of E depends (as does the general real symmetric idempotent matrix) upon infinitely many arbitrary constants. Since $(E\lambda + Q)(\lambda^2 - 1)^{-1}$ represents, as may easily be verified, for $\lambda^2 - 1 \neq 0$ the reciprocal matrix of $\lambda E - Q$ for every bounded square root Q of E, the spectrum of Q contains at most the two points $\lambda = +1$ and $\lambda = -1$. It therefore follows from

† See "Hi", p. 145.

the theorem of Hilbert † on bounded real symmetric matrices without continuous spectra that there exists for every real symmetric (and therefore bounded) square root of E an orthogonal matrix O so that $OQO^{-1} = \| \epsilon_{pq} \|$ where $\epsilon_{pq} = 0$ for $p \neq q$ and $\epsilon_{pp} = +1$ or $\epsilon_{pp} = -1$. We denote by α the number of times $+1$ occurs in the sequence $\{\epsilon_{pp}\}$ and by β the number of times -1 occurs in this sequence so that $\alpha \geqq 0$, $\beta \geqq 0$ and $\alpha + \beta = +\infty$. The numbers α and β are uniquely determined by the matrix Q, i. e. they are independent of the special choice of the matrix O employed in transforming Q into the diagonal form $\| \epsilon_{pq} \|$. It will be necessary to distinguish the following four classes of matrices which are square roots of E:

$$\text{(13)} \qquad \begin{array}{rlll} \text{I}: & \alpha = +\infty, & \beta = 2m & \text{for } 0 \leqq m < +\infty, \\ \text{II}: & \alpha = +\infty, & \beta = 2m+1 & \text{for } 0 \leqq m < +\infty, \\ \text{III}: & \alpha = m, & \beta = +\infty & \text{for } 0 \leqq m < +\infty, \\ \text{IV}: & \alpha = +\infty, & \beta = +\infty. & \end{array}$$

For the sake of brevity we shall denote an orthogonal matrix not containing -1 in its spectrum as a *Cayley matrix*. That not every orthogonal matrix is a Cayley matrix is trivial,‡ even for finite matrices; for example, no single reflection matrix (i. e. an orthogonal matrix of determinant -1) is a Cayley matrix. The justification for the name of this sub-manifold of the manifold of the orthogonal matrices is contained in the statement that the sub-manifold of the Cayley matrices may be put in a one-to-one correspondence with the manifold of the real bounded skew-symmetric matrices by means of the well known formula of Cayley,

$$\text{(14)} \qquad O = (E + S)(E - S)^{-1} = (E - S)^{-1}(E + S).$$

† See "Hi", p. 145.

‡ Cf. M. Born and P. Jordan, "Elementare Quantenmechanik" (1930), p. 37, in which they point out that the representation

$$U = (E + iH)(E - iH)^{-1} = (E - iH)^{-1}(E + iH),$$

in which H is a bounded Hermitian matrix, is restricted to those unitary matrices U not containing -1 in their spectrum. The matrix U in this representation is then and only then a real matrix, i. e. an orthogonal matrix if $H = iS$, where S is a bounded real skew-symmetric matrix, the representation becoming, in this case, the well known representation of Cayley. For finite orthogonal matrices it is not true, as is sometimes stated, [cf. for instance H. W. Turnbull, "The Theory of Determinants, Matrices and Invariants" (1928), pp. 155-157] that every orthogonal matrix of determinant $+1$ permits the Cayley representation (cf. in particular p. 588). On the other hand, those orthogonal matrices of determinant $+1$, not permitting the Cayley representation, may be regarded as limiting cases of orthogonal matrices permitting the Cayley representation. In this connection cf. L. E. Dickson, "Modern Algebraic Theories" (1926), p. 102.

In order to demonstrate this let S in (14) be a bounded real skew-symmetric matrix so that $(E+S)^{-1}$ and $(E-S)^{-1}$ exist. The matrix O in (14) is then obviously real and is orthogonal for

$$OO' = O'O = (E+S)(E-S)^{-1}(E-S)(E+S)^{-1} = E.$$

Furthermore it is a Cayley matrix for, from (14),

$$\begin{aligned} O+E &= (E+S)(E-S)^{-1}+E = (E+S)(E-S)^{-1} \\ &\quad + (E-S)(E-S)^{-1} = (E+S+E-S)(E-S)^{-1} = 2(E-S)^{-1} \end{aligned}$$

so that $(O+E)^{-1}$ exists, i. e. -1 is not a point in the spectrum of O. For the converse proposition, i. e. to any given Cayley matrix O there always exists a real bounded skew-symmetric matrix S for which (14) is valid, we consider the matrix

$$S = (O-E)(O+E)^{-1}, \tag{15}$$

which may be constructed for every Cayley matrix O. We first observe that because of

$$(O+E)(O-E) = (O-E)(O+E),$$

i. e.

$$(O-E)(O+E)^{-1} = (O+E)^{-1}(O-E),$$

we also have

$$S = (O+E)^{-1}(O-E). \tag{16}$$

Now the matrix S in (15), (16) is certainly real and bounded. In addition it is skew-symmetric, for

$$\begin{aligned} S' &= (O'+E)^{-1}(O'-E) = (O'+E)^{-1}O'O(O'-E) \\ &= [O(O'+E)]^{-1}(E-O) = (O+E)^{-1}(E-O) = -S, \end{aligned}$$

and may be employed as S in the representation (14) of O since
$(E+S)(E-S)^{-1}$

$$\begin{aligned} &= [E+(O-E)(O+E)^{-1}](O+E)(O+E)^{-1}[E-(O-E)(O+E)^{-1}]^{-1} \\ &= (O+E+O-E)(O+E-O+E)^{-1} = O. \end{aligned}$$

That this correspondence set up between the manifold of the Cayley matrices and the bounded real skew-symmetric matrices is one to one is trivial. One may see, even for finite matrices, that the manifold of the Cayley matrices is only a very restricted subset of the manifold of all orthogonal matrices. For finite matrices the manifold of the Cayley matrices is only a sub-manifold of the manifold of the rotation matrices, not even every rotation matrix being a Cayley matrix. As an example of a matrix which is a rotation matrix and

not a Cayley matrix one has the matrix $-\boldsymbol{E}$ which for even n is a rotation matrix, but is for no single value of n a Cayley matrix. This fact is emphasized very forcibly by actually attempting the representation (14) for $\boldsymbol{O}=-\boldsymbol{E}$, thereby obtaining

$$-\boldsymbol{E}=(\boldsymbol{E}+\boldsymbol{S})(\boldsymbol{E}-\boldsymbol{S})^{-1},$$

and consequently

$$-\boldsymbol{E}+\boldsymbol{S}=\boldsymbol{E}+\boldsymbol{S} \quad \text{so that} \quad -\boldsymbol{E}=\boldsymbol{E},$$

an obvious contradiction. We are accordingly convinced that, although the Cayley representation (14) possesses no convergence difficulties, it is impossible to obtain, by its use, a parametric representation of the entire manifold of orthogonal matrices. In order to obtain a parametric representation of the entire manifold of the orthogonal matrices we employ the spectral theory which will yield a canonical exponential representation.

A Hermitian matrix is said to be *reduced* if its spectrum lies in the closed interval $[0, 2\pi]$ and if its point spectrum does not contain the point 2π. According to Wintner † a one to one correspondence is set up between the manifold of the unitary matrices and the manifold of the reduced Hermitian matrices by means of the exponential representation

$$\mathrm{U}=e^{i\mathrm{H}}, \tag{17}$$

in which U is a unitary matrix and H a reduced Hermitian matrix. The definition of the matrix e^{M} is of course

$$e^{\mathrm{M}}=\sum_{\nu=0}^{\infty}(1/\nu!)\mathrm{M}^{\nu},$$

which represents for every bounded matrix M a bounded matrix. Consequently if (M, N) is a commutable pair of bounded matrices, we have

$$e^{\mathrm{M+N}}=e^{\mathrm{M}}\cdot e^{\mathrm{N}}=e^{\mathrm{N}}\cdot e^{\mathrm{M}}, \tag{18}$$

and if (M, N) is not only a commutable pair of matrices but also perpendicular [Cf. (7)]

$$e^{\mathrm{M+N}}=e^{\mathrm{M}}+e^{\mathrm{N}}-\mathrm{E}. \tag{19}$$

For our purposes it is necessary to place the manifold of the bounded Hermitian matrices in a one-to-one correspondence with the combined manifolds

† See "W", pp. 268-277. For further literature cf. H. Weyl, ***The Theory of Groups and Quantum Mechanics*** (1931), p. 399.

of the bounded real symmetric and real skew-symmetric matrices by means of the linear representation

$$(20) \qquad H = \pi B + iS, \quad B = \bar{B} = B', \quad S = \bar{S} = -S'.$$

In Theorem I we show that from a consideration of a trigonometric momentum problem with real "momenta" the entire manifold of the orthogonal matrices may, by means of the exponential representation (17), (20), i. e.

$$(21) \qquad O = e^{iH} = e^{i[\pi B + iS]},$$

be put in a one to one correspondence † with that sub-manifold of the reduced Hermitian matrices for which in (20)

$$(22) \qquad B^2 = B, \quad BS = SB,$$

i. e. B is a real symmetric idempotent (and therefore bounded) matrix and S a bounded real skew-symmetric matrix, commutable with B, and such that $\pi B + iS$ is a reduced Hermitian matrix.

On p. 606 we introduce the term *reduced skew-symmetric matrix* for those skew-symmetric matrices S, the entire spectrum of which is contained in the closed interval $[-i\pi, +i\pi]$ of the imaginary axis, and whose point spectrum does not contain the end-points $-i\pi$ and $+i\pi$. We then show in Lemma I that the manifold of the reduced Hermitian matrices (20), (22) may be put in a one-to-one correspondence with the manifold of the perpendicular matrix pairs $(E-B, S)$ in which S is a reduced skew-symmetric matrix. As a consequence of Lemma I we then obtain Theorem II which states that the entire manifold of the orthogonal matrices may, by means of the exponential representation (21), be put in a one-to-one correspondence with the manifold of the matrix pairs (B, S) where

$$(23) \qquad \begin{gathered} B = \bar{B} = B',\ B^2 = B; \quad S = \bar{S} = -S'; \\ (E-B)S = S(E-B) = \|0\|, \quad S \text{ is reduced.} \end{gathered}$$

We then show in Theorem III that the one-to-one representation (21), (22) is in reality none other than the one-to-one representation

$$(24) \qquad O = 2(E-B) - e^{-S}$$

in which the matrix pair (B, S) is identical with the matrix pair (B, S) in (21), (22).

In Lemma II on p. 615, we show that the manifold of the matrices Q which are square roots of E may be placed in a one-to-one correspondence

† See "W", p. 277.

with the manifold of the idempotent matrices I by means of the linear representation

$$Q = E - 2I. \tag{25}$$

An application of this lemma to Theorem I then yields Theorem IV on p. 615, in which the entire manifold of the orthogonal matrices is shown to be in a one-to-one correspondence, by means of the representation

$$O = Q \cdot e^{-S} = e^{-S} \cdot Q \tag{26}$$

with the manifold of the matrix pairs (Q, S) in which Q is a real symmetric square root of E (and therefore itself an orthogonal matrix) and S a bounded real skew-symmetric matrix, commutable with Q, and such that the Hermitian matrix $(\pi/2)(E - Q) + iS$ is reduced.

Since Theorem IV assigns to every orthogonal matrix O a uniquely determined real symmetric square root Q of E we may interpret the classification (13) of the real symmetric square roots of E as a classification of the orthogonal matrices into the four classes of p. 617. As immediate examples of the four types of orthogonal matrices we have the matrices Q of (124) themselves.

For normal forms, under orthogonal matrix transformations of the matrices of each of the four classes we obtain, as consequences of Theorems I and IV and Lemma I, the canonical forms given in (133) in which the matrices $Q^{(i)}$ and $S^{(i)}$ $(i = 1, 2, 3, 4)$ are defined in (124) and (125) respectively. The matrices $\boldsymbol{O}_{2m}$ and $\boldsymbol{O}_{2m+1}$ occurring in I and II of (133) are the canonical finite rotation matrices as defined in (131). From these canonical forms I and II of (133) it is clear at once that the orthogonal matrices of Classes I and II can be interpreted, with respect to a suitably orientated coördinate system, as rotations and reflections (in the geometrical sense of the word) in some subspace of the real Hilbert space, possessing a sufficiently high but finite number of dimensions, in which a subspace of infinitely many dimensions is not "affected" † by the transformation. The orthogonal matrices of Class III may accordingly be interpreted as representing those orthogonal transformations of the real Hilbert space which leave a subspace of at most a finite number of dimensions "unaffected" and the orthogonal matrices of Class IV as the orthogonal transformations of the real Hilbert space which

† For example we say that the transformation in the ordinary Euclidean 3-space possessing the matrix

$$\begin{Vmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 0 & 0 & 1 \end{Vmatrix}$$

leaves a space of one dimension not "affected" by the transformation.

leave "unaffected" a subspace of infinitely many dimensions while simultaneously "affecting" a complement subspace of infinitely many dimensions. It may not be superfluous to remark that according to this classification the unit matrix E is an orthogonal matrix of Class I and the orthogonal matrix $-E$ of Class III. It is still an open question as to whether the orthogonal matrices of Classes III and IV may be divided into "rotations" and "reflections." In this connection cf. p. 592, where we are able to give a precise formulation of this unsolved problem.

During the last few years several authors have investigated rotations of the functional space, or what is the same, rotations in the Hilbert space. These papers, however, deal only with very special orthogonal matrices, supposing, for instance, that the determinant of the matrix is convergent and that the "deviation" of the orthogonal matrix from the unit matrix is a completely continuous (vollstetig in the sense of Hilbert) matrix. Even these very restrictive conditions are not sufficient for the validity of the determinant apparatus as employed by them. In this paper we deal with the totality of the orthogonal transformations of the Hilbert space and are, therefore, denied any recourse to the apparatus of determinants. For example the inaptness of the determinant apparatus is clearly revealed by attempting to differentiate, by its use, between the various classes (133) of the orthogonal matrices $Q^{(i)}$ of (124).

The Classes I and II of the orthogonal matrices O are then shown to be the natural extensions of the finite rotation and reflection matrices as defined on p. 579. The matrices of Classes III and IV obviously have no analogues for finite matrices. At this point there is introduced a definition which divides † the entire manifold of the infinite orthogonal matrices into two sub-manifolds in such a manner that the definition when applied to the manifold of the finite orthogonal matrices divides it into two sub-manifolds which are identical with the sub-manifolds of the finite rotation and reflection matrices as defined on p. 579. We define on p. 620 that sub-manifold of the orthogonal matrices, the elements O of which may be represented in the form e^{G} where G is a real bounded skew-symmetric matrix [not necessarily uniquely determined by O (Cf. p. 594)], i. e.

$$(27) \qquad O = e^{G}, \qquad G = \bar{G} = -G',$$

as the manifold of the *rotation matrices* and the complement of this manifold with respect to the entire manifold of the orthogonal matrices [i. e. the manifold of those matrices not permitting the exponential representation (27)],

† Cf. p. 579.

as the manifold of the *reflection matrices*. This definition of the manifolds of the rotation and reflection matrices will be referred to as the exponential definition in order to distinguish it from the determinant definition on p. 579 which is utterly valueless for infinite orthogonal matrices since the determinant of an orthogonal matrix is convergent only for very special orthogonal matrices. In order to justify this definition we show on p. 620 in Lemma III that the exponential definition of the rotation and reflection matrices is, for finite matrices, identical with the determinant definition. In Theorem V on p. 621, we show that the matrices of Classes I and II are, in the sense of the exponential definition, respectively rotation and reflection matrices. The first part of Theorem V follows directly from the canonical form for the matrices of Class I. In order to prove the second part of Theorem V we show in Lemma IV on p. 622 that if θ (θ real) is a point in the spectrum of an arbitrary (not necessarily reduced) Hermitian matrix H then $e^{i\theta}$ is a point in the spectrum of e^{iH}. As a corollary to Lemma IV we have on p. 624 that if $i\theta$ is a point in the spectrum of an arbitrary real bounded skew-symmetric matrix S then $e^{i\theta}$ is a point in the spectrum of e^{S} from which, and the canonical form II of (133), the second part of Theorem V is shown to follow.

The unsolved problem, mentioned on p. 591, as to whether the orthogonal matrices of Classes III and IV permit a resolution into "*rotations*" and "*reflections*" may now be put in a precise form as follows: Do the Classes III and IV contain both rotation and reflection matrices as defined in (27)? That there exist rotation matrices in the sense of the exponential definition (27) is trivial since the orthogonal matrices $Q^{(3)}$ and $Q^{(4)}$ of (124) are of Classes III and IV respectively and obviously permit the exponential representation (27). The unsolved part of the problem then requires the demonstration of the existence or the non-existence of reflection matrices, as defined in (27), for each of the Classes III and IV. Put more directly, do the matrices of the Classes III and IV all permit the exponential representation (27)? Since the property of a matrix to be a rotation matrix, in the sense of the exponential definition (27), is obviously invariant † with respect to orthogonal transformations, it is sufficient to treat this question for the canonical forms III and IV of (133) for the matrices of Classes III and IV. From the fact that the orthogonal matrices $Q^{(3)}$ and $Q^{(4)}$ in the canonical forms (133) are both rotation matrices, it is trivial that if the

† This invariance property is trivial inasmuch as follows from the definition of e^{G} on p. 13 that if O is an arbitrary orthogonal matrix then $Oe^{G}O^{-1} = e^{OGO^{-1}}$ in which OGO^{-1} is again a bounded real skew-symmetric matrix.

rotation matrices [i. e. those orthogonal matrices permitting the representation (27)] form a group, the canonical matrices III and IV of (133) are necessarily rotation matrices [i. e. also permit the representation (27)]. It may not be superfluous to point out that, because of the special character of the matrices occurring in the canonical matrices III and IV of (133), the condition that the rotation matrices (27) form a group is not a necessary condition for the matrices of Classes III and IV to be rotation matrices (27). A treatment of the above unsolved problems seem to require a detailed investigation of orthogonal and real skew-symmetric matrices possessing continuous spectra. Connected with this problem is the question † as to whether the manifold of all orthogonal matrices forms a connected manifold or not (with respect to some natural metric in the manifold of all matrices). For the orthogonal transformations of a space with a finite number of dimensions the manifold of all orthogonal matrices is not connected (with respect to every natural metric of the manifold) but is divided into two separate manifolds (Cf. p. 579). In other words the Euclidean space of a finite number of dimensions may be said to be an "orientable" space. It has not been possible as yet to answer the question, on the basis of the parametrical representation of orthogonal matrices given in the paper, as to whether the real Hilbert space forms, in this sense, an orientable space or not. In order that the manifold of all orthogonal matrices be connected it is necessary that each of the sub-manifolds in (133) be connected. It seems likely that for the answer to the above question the extensive use of the theory of the spectral matrix of the orthogonal matrices is unavoidable, the problems treated in the present paper depending upon the theory of continuous spectra of unitary matrices (cf. "W", pp. 257-258).

Theorems I, II, III and IV all give a complete parametric representation of the manifold of the orthogonal matrices and we saw that (27) where G is an arbitrary real bounded skew-symmetric matrix, may be defined as the parametric representation of the manifold of the rotation matrices of the Hilbert space. This parametric representation of the rotation matrices is the substitute † in Hilbert space for the Euler-Hurwitz factorization ‡ into products of "*two-block*" rotation matrices which breaks down for infinite

† See "W", p. 277.

‡ L. Euler, *Commentationes arithmeticae collectae*, Vol. 1, p. 427 and A. Hurwitz, "Ueber die Erzeugung der Invarianten durch Integration," *Göttingen Nachrichten* (1897), pp. 76-77. In connection with the work of Hurwitz, cf. I. Schur, "Neue Anwendungen der Integralrechnung auf Probleme der Invariantentheorie," *Sitzungsberichte der Preussischen Akademie der Wissenschaften* (1924), p. 196, where a statement of Hurwitz is corrected.

matrices since there exists no last dimension. There are even other obstacles to a generalization of the Euler-Hurwitz representation inasmuch as the following example shows that a matrix which is (even in the sense of a "*weak*" convergence) the limit of a sequence of orthogonal "*two block*" matrices is not necessarily itself an orthogonal matrix. Denote by E_{pq} the "*two- block*" orthogonal matrix obtained from the infinite unit matrix E by interchanging the p-th and q-th columns (or what amounts to the same thing, by interchanging the p-th and q-th rows). Consider the sequence of orthogonal matrices $\{O_i\}$ defined by $O_1 = E$, $O_i = O_{i-1}E_{i-1\,i}$ $(i = 2, 3, \cdots)$, i. e. the sequence

$$O_1 = E, \quad O_2 = E_{12}, \quad O_3 = E_{12}E_{23}, \cdots,$$

for which

$$\lim_{n\to+\infty} O_n = \begin{Vmatrix} 0 & 0 & 0 & 0 & . & . & . \\ 1 & 0 & 0 & 0 & . & . & . \\ 0 & 1 & 0 & 0 & . & . & . \\ 0 & 0 & 1 & 0 & . & . & . \\ . & . & . & . & & & \\ . & . & . & . & & & \\ . & . & . & . & & & \end{Vmatrix},$$

and which is certainly not an orthogonal matrix.

It may be mentioned that to a given rotation matrix O there may exist infinitely many real bounded skew-symmetric matrices which may be employed as the matrix G in (27). This is illustrated, even for finite matrices, by the example

$$\begin{Vmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{Vmatrix} = \exp \begin{Vmatrix} 0 & (\theta + 2k\pi) \\ -(\theta + 2k\pi) & 0 \end{Vmatrix}; \quad (k = 0, \pm 1, \cdots).$$

It would be desirable to obtain a simple sub-manifold of the matrices G in (27) so that the matrices of this sub-manifold are in one-to-one correspondence with the rotation matrices. Certain considerations for finite matrices seem to indicate that this may be achieved for finite matrices by restricting the representation (27) to a sub-manifold analogous to the manifold of the reduced skew-symmetric matrices, and it is hoped to treat this question among others in a later appearing paper on the finite matrices.

For finite matrices it is possible † to obtain, for $n = 2$ and $n = 3$, the formulas of Rodrigues from the exponential representation (27).

† Cf. G. Galanti, "Algoritmi di calcolo motoriale," *Atti Accademia Nazionale dei Lincei*, Vol. 6, series 13 (1931), pp. 861-866.

Finally, it may be mentioned that the methods and results of the present paper hold, precisely as the spectral theory of unitary transformations or the method of trigonometric momentum problem used in this paper, not only for matrices but also for operators in the Hilbert space. However, since the possibility of such generalizations is known to be trivial we prefer, for clearness, to deal with matrices only.

I. Reality Discussion of the Trigonometrical Moment Problem of Herglotz.

We consider the trigonometrical moment problem †

$$(28) \qquad 2\pi c_n = \int_0^{2\pi} e^{in\phi}\, d\tau(\phi)\,; \qquad (n = 0, \pm 1, \cdots),$$

together with its homogeneous special case

$$(28') \qquad 0 = \int_0^{2\pi} e^{in\phi}\, d\tau(\phi)\,; \qquad (n = 0, \pm 1, \cdots),$$

in which the left-hand members are preassigned real or complex numbers and $\tau(\phi)$ an unknown real or complex function of the real argument ϕ whose definition is desired in the closed interval $0 \leqq \phi \leqq 2\pi$. In the following we shall consider only real solutions $\tau(\phi)$ and therefore suppose that

$$(29) \qquad c_n = \bar{c}_{-n}\,; \qquad (n = 0, 1, \cdots).$$

It will be convenient to employ the notations

$$(30) \qquad \begin{aligned} &\underset{\beta-0}{\Delta} f(\phi) = f(\beta) - f(\beta - 0), \\ &\underset{\beta}{\Delta} f(\phi) = f(\beta + 0) - f(\beta - 0), \\ &\underset{\beta+0}{\Delta} f(\phi) = f(\beta + 0) - f(\beta), \end{aligned}$$

(provided these limits exist) for the various "springs" of an arbitrary real or complex function $f = f(\phi)$ at the point $\phi = \beta$ and the customary notations $[a, b]$ and (a, b) for the closed and open intervals $a \leqq \phi \leqq b$ and $a < \phi < b$ respectively.

We suppose that the moment problem (28) permits a solution $\tau(\phi)$ of bounded variation in the interval of its definition $[0, 2\pi]$ so that the limits

† G. Herglotz, "Über Potenzreihen mit positiven, reellen Teil im Einheitskreis," *Königlich Sächsische Gesellschaft der Wissenschaften zu Leipzig, Mathematisch-Physische Klasse* (1911), p. 501-511. This paper will be referred to as "He".

(31) $\tau(\phi - 0)$ for $0 < \phi \leqq 2\pi$; $\tau(\phi + 0)$ for $0 \leqq \phi < 2\pi$,

exist. It is known † from the general theory of Stieltjes integrals that if (28) possesses a solution $\tau(\phi)$ which is of bounded variation in the interval of its definition $[0, 2\pi]$ it is uniquely determined, up to an additive constant, on the set of its continuity points contained in the interval $(0, 2\pi)$, its values at its inner discontinuity points (i. e. for $0 < \phi < 2\pi$) being assignable at will. Since the set of continuity points of $\tau(\phi)$ in the interval $(0, 2\pi)$ is necessarily everywhere dense in this interval the values of $\tau(0+0)$ and $\tau(2\pi - 0)$ may then be calculated [cf. (31)] from the values of $\tau(\phi)$ at its continuity points and the values of $\tau(0)$ and $\tau(2\pi)$ chosen at will, provided that the conditions

$$(32)\quad \underset{0+0}{\Delta}\tau(\phi) + \underset{2\pi-0}{\Delta}\tau(\phi) = 2\pi c_0 - \int_{0+0}^{2\pi-0} d\tau(\phi) = 2\pi c_n - \int_{0+0}^{2\pi-0} e^{in\phi}\, d\tau(\phi); \qquad (n = 0, \pm 1, \cdots),$$

which arise from (28), are satisfied.‡ The only solutions of bounded variation of the homogeneous moment problem (28′) may in turn be completely characterized by the properties §

$$(33)\quad \tau(\phi) = \alpha = \bar{\alpha} = \text{const. at the continuity points of } \tau(\phi) \text{ in } (0, 2\pi),$$

$$(34)\quad \underset{0+0}{\Delta}\tau(\phi) + \underset{2\pi-0}{\Delta}\tau(\phi) = 0,$$

inasmuch as, in (32), we have $c_0 = 0$ and $\int_{0+0}^{2\pi-0} d\tau(\phi) = 0$.

A further possible specialization of the character of the solution $\tau(\phi)$ of (28) is that it be monotone non-decreasing. The necessary and sufficient conditions on the c_n in order that the moment problem (28) possess a monotone non-decreasing solution $\tau(\phi)$ are given by Herglotz ¶ and state that the truncated matrices (Abschnittsmatrizen) of the infinite Hermitian matrix $\| c_{\mu-\nu} \|$, introduced by Toeplitz, be not negative definite, i. e.

$$(35)\quad \sum_{\mu=1}^{n}\sum_{\nu=1}^{n} c_{\mu-\nu}\xi_\mu\bar{\xi}_\nu \geqq 0 \text{ for arbitrary } \xi_1, \xi_2, \cdots; \qquad (n = 1, 2, \cdots),$$

where $c_{\mu-\nu} = c_k = \bar{c}_{-k}$ for $k = \mu - \nu$.

---

† Cf. F. Hausdorff, "Über das Momentenproblem für ein endliches Interval," *Mathematische Zeitschrift*, Vol. 16 (1923), pp. 220-248. This paper will be referred to as "Ha". Cf. also O. Perron, "Die Lehre von den Kettenbrüchen" (1913), pp. 362-374. This book will be referred to as "P".

‡ See "W", p. 268.

§ See "Ha", p. 241.

¶ See "He".

We denote by $E = \| \delta_{pq} \|$ the unit matrix and by $\Phi[A; x]$ the compound form (Kopplungsform)

$$(36) \qquad \Phi[A; x] = \sum_{p=1}^{+\infty} \sum_{q=1}^{+\infty} a_{pq} x_p \bar{x}_q,$$

of an arbitrary bounded matrix $A = \| a_{pq} \|$. With these notations we consider the trigonometric moment problem † arising from (28) if one places

$$(37) \qquad c_n = \Phi[U^n; x]; \qquad (n = 0, \pm 1, \cdots),$$

i. e., the moment problem

$$(38) \quad 2\pi\Phi[U^n; x] = \int_0^{2\pi} e^{in\phi}\, d\tau(\phi; x), \qquad |x| = 1; \; (n = 0, \pm 1, \cdots),$$

where U is a given unitary matrix so that for every unit vector x [i. e., $|x| = 1$] there belongs the sequence of momenta c_k defined in (37) satisfying the conditions (29). It is known that this moment problem always possesses a real solution $\tau(\phi; x)$ which depends upon the vector x and is, as a matter of fact, a bounded Hermitian form

$$(39) \qquad \tau(\phi; x) \equiv \Phi[\| \tau_{pq}(\phi) \|; x],$$

where

$$\tau_{pq}(\phi) = \bar{\tau}_{qp}(\phi), \qquad 0 \leqq \phi \leqq 2\pi,$$

and the indices p, q take, as will always be understood in the remainder of the paper, unless explicitly stated to the contrary, independently the values $1, 2, \cdots$. The Hermitian form (39) is then, for an arbitrary given unit vector x, uniquely determined up to an additive constant at all of its continuity points in the interval $(0, 2\pi)$. Moreover, since the conditions (35) are known to be fulfilled by the momenta (37), for every unit vector x, the Hermitian form (39) may be chosen to be, for an arbitrary given unit vector x, a monotone non-decreasing function of ϕ in the interval $(0, 2\pi)$. The values $\tau(0, x)$ and $\tau(2\pi, x)$ of the Hermitian form at the end points may then, for an arbitrary given unit vector x, be assigned at will subject to the condition that $\tau(\phi; x)$ be monotone in the closed interval $[0, 2\pi]$ and that

$$(40) \qquad \underset{0+0}{\Delta} \tau(\phi; x) + \underset{2\pi-0}{\Delta} \tau(\phi; x) = 2\pi - \int_{0+0}^{2\pi-0} d\tau(\phi; x).$$

We now remove the arbitrary character of the solution $\tau(\phi; x)$ by prescribing, for an arbitrary given unit vector x, the following normalizing conditions:

---

† See "W", p. 269.

$$(41)\qquad \tau(0;x)\equiv 0,$$

$$(42)\qquad \tau(\phi-0;x)\equiv\tau(\phi;x)\quad\text{for}\quad 0<\phi<2\pi,$$

$$(43)\qquad \underset{2\pi-0}{\Delta}\tau(\phi;x)\equiv 0.$$

The normalizing condition (42) serves to fix the value of the solution $\tau(\phi;x)$ at its discontinuity points in the interval $(0,2\pi)$. The prescribing of the normalizing condition (43) then from (40) serves to determine uniquely the "spring" $\underset{0+0}{\Delta}\tau(\phi;x)$. Finally the normalizing condition (41) fixes the value of the additive constant so that the solution $\tau(\phi;x)$ is uniquely determined at all points of the closed interval $[0,2\pi]$ since, for $\phi=2\pi$, we have from (41) and (38), for $n=0$,

$$(44)\qquad \tau(2\pi;x)\equiv 2\pi,\qquad |x|=1.$$

The matrix of the Hermitian form under the normalizing conditions (41), (42) and (43) has been introduced by Wintner † and designated by him as the spectral matrix of U. Since (38), (41), (42), (43) and (44) hold for an arbitrary unit vector x the elements of the spectral matrix $\|\tau_{pq}(\phi)\|$ obviously possess the following properties ‡:

$$(38')\qquad 2\pi U^n=\left\|\int_0^{2\pi}e^{in\phi}\,d\tau_{pq}(\phi)\right\|;\qquad (n=0,\pm1,\cdots),$$

$$(41')\qquad \tau_{pq}(0)=0,$$

$$(42')\qquad \tau_{pq}(\phi-0)=\tau_{pq}(\phi)\quad\text{for}\quad 0<\phi<2\pi,$$

$$(43')\qquad \underset{2\pi-0}{\Delta}\tau_{pq}(\phi)=0\quad\text{i. e.}\quad \tau_{pq}(2\pi-0)=\tau_{pq}(2\pi),$$

$$(44')\qquad \tau_{pq}(2\pi)=2\pi\delta_{pq}.$$

If now the unitary matrix is assumed to be real, i. e. U is an orthogonal matrix $U=\bar{U}=O$, the elements of the spectral matrix possess in addition to the above mentioned properties the following two properties:

I. *In order that the elements $\tau_{pq}(\phi)$ of the spectral matrix of the unitary matrix U be the elements of the spectral matrix of a real unitary matrix $U=\bar{U}$, i. e. an orthogonal matrix O, it is necessary and sufficient that*

$$(45)\qquad \omega_{pq}(\phi)=b\delta_{pq}$$

at the continuity points of $\omega_{pq}(\phi)$ in $(0,2\pi)$, where

† See "W", p. 269.

‡ Cf. the uniqueness property of compound forms mentioned on p. 580.

(46) $\qquad \omega_{pq}(\phi) = \bar{\tau}_{pq}(\phi) + \tau_{pq}(2\pi - \phi) \quad$ *and* $\quad b = \bar{b} =$ *const.*

II. *The "springs"* $\underset{0+0}{\Delta}\, \tau_{pq}(\phi)$ *of the elements of the spectral matrix of an orthogonal matrix* O *are real.*

The demonstration of the two statements I and II proceeds simultaneously and we begin by showing the necessity of (45). We accordingly assume $\mathrm{U} = \bar{\mathrm{U}} = 0$ so that the conjugate equations of (38′) become

$$2\pi \mathrm{O}^n = \| \int_0^{2\pi} e^{-in\phi}\, d\bar{\tau}_{pq}(\phi) \|; \qquad (n = 0, \pm 1, \cdots). \tag{47}$$

On the other hand if we introduce a new integration variable θ in (38′) by means of the transformation

$$\phi = 2\pi - \theta, \tag{48}$$

we obtain

$$-2\pi \mathrm{O}^n = -\| \int_{2\pi}^{0} e^{in(2\pi-\theta)}\, d\tau_{pq}(2\pi - \theta) \| = \| \int_0^{2\pi} e^{-in\theta}\, d\tau_{pq}(2\pi - \theta) \|. \tag{49}$$

$$(n = 0, \pm 1, \cdots).$$

On adding (47) and (49) we have from (46)

$$\| 0 \| = \| \int_0^{2\pi} e^{-in\phi}\, d\omega_{pq}(\phi) \|; \qquad (n = 0, \pm 1, \cdots), \tag{50}$$

so that the Hermitian form $\omega(\phi; x) = \Phi\,[\| \omega_{pq}(\phi) \|; x]$ is, for any arbitrary given unit vector x, a solution of the homogeneous moment problem (28′). Furthermore the Hermitian form $\Phi\,[\| \tau_{pq}(\phi) \|; x]$ and consequently the Hermitian form $\Phi\,[\| \bar{\tau}_{pq}(\phi) \|; x]$ are, independently of the unit vector x, monotone non-decreasing functions of ϕ in the interval $[0, 2\pi]$. The Hermitian form $\Phi\,[\| \tau_{pq}(2\pi - \phi) \|; x]$ is, on the contrary, monotone non-increasing in $[0, 2\pi]$, for any given unit vector x. Accordingly the Hermitian form $\omega(\phi; x)$ is of bounded variation [cf. (46)] in the interval $[0, 2\pi]$ for any arbitrary given unit vector x. From the uniqueness property (33) of solutions of bounded variation of the homogeneous moment problem (28′) we have, for an arbitrary given unit vector x,

$$\omega(\phi; x) \equiv \Phi\,[\| \bar{\tau}_{pq}(\phi) + \tau_{pq}(2\pi - \phi) \|; x] = b = \text{const.}, \tag{51}$$

at the continuity points of $\omega_{pq}(\phi)$ in $(0, 2\pi)$, and consequently, since (51) holds for an arbitrary unit vector x, the necessity of the condition (45) of I. Furthermore, since $\| \tau_{pq}(\phi) \|$ is Hermitian, the constant b in (51) is real.

The statement II is an immediate consequence of the necessary condition

(45) and the normalizing conditions (41′), (43′), (44′). From (45) we obtain on observing (41′)

$$(52) \qquad \bar{\tau}_{pq}(\phi) - \bar{\tau}_{pq}(0) = b\delta_{pq} - \tau_{pq}(2\pi - \phi)$$

at the continuity points of $\omega_{pq}(\phi)$ in $(0, 2\pi)$. Since the continuity points of $\omega_{pq}(\phi)$, which is obviously of bounded variation in $[0, 2\pi]$, are certainly everywhere dense in the interval $(0, 2\pi)$ we have for $\phi \to +0$ [cf. (30)] from (43′) and (44′)

$$(53) \qquad \underset{0+0}{\Delta}\, \bar{\tau}_{pq}(\phi) = (b - 2\pi)\delta_{pq} \quad \text{where} \quad b = \bar{b} = \text{const.},$$

i. e. II is proven.

We now show the sufficiency of the condition (45) in I. From (38′) for $n = 1$ we have as a consequence of (43′)

$$2\pi U = \| \int_0^{2\pi} e^{i\phi}\, d\tau_{pq}(\phi) \| = \| \underset{0+0}{\Delta}\, \tau_{pq}(\phi) \|$$
$$+ \| \int_{0+0}^{\pi-0} e^{i\phi}\, d\tau_{pq}(\phi) \| - \| \underset{\pi}{\Delta}\, \tau_{pq}(\phi) \| + \| \int_{\pi+0}^{2\pi-0} e^{i\phi}\, d\tau_{pq}(\phi) \|,$$

which may, on introducing a new integration variable θ by means of (48) in $\int_{\pi+0}^{2\pi-0}$, be written in the form

$$(54) \qquad 2\pi U = \| \underset{0+0}{\Delta}\, \tau_{pq}(\phi) \| + \| \int_{0+0}^{\pi-0} e^{i\phi}\, d\tau_{pq}(\phi) \|$$
$$- \| \underset{\pi}{\Delta}\, \tau_{pq}(\phi) \| - \| \int_{0+0}^{\pi-0} e^{-i\theta}\, d\tau_{pq}(2\pi - \theta) \|.$$

We now utilize (45) to obtain

$$(55) \qquad \int_{0+0}^{\pi-0} e^{-i\phi}\, d\omega_{pq}(\phi) = \int_{0+0}^{\pi-0} e^{-i\phi}\, d\bar{\tau}_{pq}(\phi) + \int_{0+0}^{\pi-0} e^{-i\phi}\, d\tau_{pq}(2\pi - \phi) = 0,$$

inasmuch as the values of $\omega_{pq}(\phi)$ at the inner discontinuity points are immaterial in the Stieltjes integration. From (55) we may write (54) in the form

$$2\pi U = \| \underset{0+0}{\Delta}\, \tau_{pq}(\phi) \| + \| \int_{0+0}^{\pi-0} e^{i\phi} d\tau_{pq}(\phi) + \int_{0+0}^{\pi-0} e^{-i\phi} d\bar{\tau}_{pq}(\phi) \| - \| \underset{\pi}{\Delta}\, \tau_{pq}(\phi) \|,$$

so that from II there only remains to show that $\| \underset{\pi}{\Delta}\, \tau_{pq}(\phi) \|$ is a real matrix. This follows immediately on writing (45) in the form

$$\tau_{pq}(2\pi - \phi) - \tau_{pq}(\phi) = b\delta_{pq} - [\tau_{pq}(\phi) + \bar{\tau}_{pq}(\phi)],$$

valid for the continuity points of $\omega_{pq}(\phi)$ in $(0, 2\pi)$, and then taking the limit for $\phi \to \pi - 0$.

II. Application of the Trigonometrical Moment Problem to Orthogonal Matrices.

Among the properties of the spectral matrix $\| \tau_{pq}(\phi) \|$ of a given unitary matrix U, Wintner † has shown that it is possible to employ it in the construction of a bounded Hermitian matrix H

$$(56) \qquad H = \| (1/2\pi) \int_0^{2\pi} \phi d\tau_{pq}(\phi) \|,$$

the elements of the spectral matrix $\| \sigma_{pq}(\phi) \|$ of which are given by

$$(57) \qquad \begin{aligned} \sigma_{pq}(\phi) &= 0 && \text{for} \quad -\infty < \phi \leqq 0, \\ 2\pi\sigma_{pq}(\phi) &= \tau_{pq}(\phi) && \text{for} \quad 0 \leqq \phi \leqq 2\pi, \\ \sigma_{pq}(\phi) &= \delta_{pq} && \text{for} \quad 2\pi \leqq \phi < +\infty. \end{aligned}$$

Conversely if H is a bounded Hermitian matrix whose spectral matrix $\| \sigma_{pq}(\phi) \|$ fulfills the conditions

$$(58) \qquad \begin{aligned} \sigma_{pq}(\phi) &= 0 && \text{for} \quad -\infty < \phi \leqq 0, \\ \sigma_{pq}(2\pi - 0) = \sigma_{pq}(\phi) &= \delta_{pq} && \text{for} \quad 2\pi \leqq \phi < +\infty, \end{aligned}$$

(i. e. $\phi = 2\pi$ is not a point of the point spectrum of H) and if a matrix $\| \tau_{pq}(\phi) \|$ be introduced by means of the definitions

$$(59) \qquad \tau_{pq}(\phi) = 2\pi\sigma_{pq}(\phi) \quad \text{for} \quad 0 \leqq \phi \leqq 2\pi,$$

then the matrix

$$(60) \qquad U = e^{iH} = \| (1/2\pi) \int_0^{2\pi} e^{i\phi}\, d\tau_{pq}(\phi) \|,$$

is a unitary matrix possessing as its spectral matrix the matrix $\| \tau_{pq}(\phi) \|$. If we designate a Hermitian matrix whose spectrum lies in the interval $[0, 2\pi]$ and whose point spectrum does not contain the point 2π, as a reduced Hermitian matrix we may sum up the above results in the following form: There exists between the unitary matrices and the reduced Hermitian matrices a one-to-one correspondence, i. e. to a given unitary matrix U there is associated a reduced Hermitian matrix H by means of (56) and to this same reduced Hermitian matrix H there is associated, by means of (60), the previously given unitary matrix U.

In the following we propose to apply the results of the previous section in order to characterize that sub-class of the reduced Hermitian matrices which is in one-to-one correspondence with the real unitary matrices, i. e.

† See "W", pp. 268-277.

12

the orthogonal matrices. In any case we may write any Hermitian matrix uniquely in the form

$$(61) \qquad \mathrm{H} = \pi\mathrm{B} + i\mathrm{S},$$

where B is a real symmetric matrix and S a real skew-symmetric matrix so that our problem † is to characterize the matrix pair (B, S) for which the reduced Hermitian matrix $\mathrm{H} = \pi\mathrm{B} + i\mathrm{S}$ has the property that the matrix $\mathrm{U} = e^{i\mathrm{H}}$ is real. As an answer to this problem we show

THEOREM I. *The necessary and sufficient conditions in order that the reduced Hermitian matrix* $\mathrm{H} = \pi\mathrm{B} + i\mathrm{S}$ *yield a real unitary matrix* O, *i. e. an orthogonal matrix, in the one-to-one representation*

$$(62) \qquad \mathrm{O} = e^{i\mathrm{H}} = e^{i(\pi\mathrm{B}+i\mathrm{S})},$$

are that S *be a bounded real skew-symmetric matrix and* B *a real symmetric idempotent matrix which is commutable with* S, *i. e.*

$$(63a) \qquad \mathrm{B} = \bar{\mathrm{B}} = \mathrm{B}', \quad \mathrm{B}^2 = \mathrm{B}; \quad \mathrm{S} = \bar{\mathrm{S}} = -\mathrm{S}',$$

$$(63b) \qquad \mathrm{BS} = \mathrm{SB}.$$

Before beginning the demonstration of Theorem I we note that the representation (62) may, from (63b) and the distributive property (18) of the exponential with respect to commutable matrices, be given the form

$$(62') \qquad \mathrm{O} = e^{i\pi\mathrm{B}} \cdot e^{-\mathrm{S}} = e^{-\mathrm{S}} \cdot e^{i\pi\mathrm{B}}.$$

It may also be mentioned that the representation (62) or its equivalent (62') is "covariant" under orthogonal matrix transformation. More precisely if V is any orthogonal matrix then we have from (62)

$$\mathrm{VOV}^{-1} = \exp\,[i\pi(\mathrm{VBV}^{-1} + i\mathrm{VSV}^{-1})],$$

where VBV^{-1} is a real symmetric idempotent matrix [cf. p. 585] commutable with the bounded real skew-symmetric matrix VSV^{-1} and $\pi\mathrm{VBV}^{-1} + i\mathrm{VSV}^{-1}$ is a reduced Hermitian matrix if $\pi\mathrm{B} + i\mathrm{S}$ is a reduced Hermitian matrix. This "covariance" property is important in what follows inasmuch as without it the classification of the orthogonal matrices in four classes as developed below would be meaningless.

We begin with a demonstration of the necessity of the conditions (63a) and (63b) and shall retain the notation $\| \tau_{pq}(\phi) \|$ for the spectral matrix of a unitary matrix U. First we show that,

† See "W", p. 277.

$$(64)\qquad \int_{0+0}^{\pi-0} d\bar{\tau}_{pq}(\phi) = \int_{\pi+0}^{2\pi-0} d\tau_{pq}(\phi),$$

$$(65)\qquad \int_{\pi+0}^{2\pi-0} \phi d\tau_{pq}(\phi) = 2\pi \int_{0+0}^{\pi-0} d\bar{\tau}_{pq}(\phi) - \int_{0+0}^{\pi-0} \phi d\bar{\tau}_{pq}(\phi).$$

On introducing a new integration variable θ by means of (48) we obtain for the left-hand members of (64) and (65) respectively

$$(66)\qquad \int_{0+0}^{\pi-0} d\bar{\tau}_{pq}(\phi) = \int_{2\pi-0}^{\pi+0} d\bar{\tau}_{pq}(2\pi-\theta).$$

$$(67)\qquad \int_{\pi+0}^{2\pi-0} \phi d\tau_{pq}(\phi) = \int_{\pi-0}^{0+0} (2\pi-\theta)\, d\tau_{pq}(2\pi-\theta).$$

From (45) we obtain

$$(68)\qquad \int_{2\pi-0}^{\pi+0} d\omega_{pq}(\phi) = \int_{2\pi-0}^{\pi+0} d\bar{\tau}_{pq}(\phi) + \int_{2\pi-0}^{\pi+0} d\tau_{pq}(2\pi-\phi) = 0,$$

$$(69)\qquad \int_{\pi-0}^{0+0} (2\pi-\phi)\, d\omega_{pq}(\phi) = \int_{\pi-0}^{0+0} (2\pi-\phi)\, d\bar{\tau}_{pq}(\phi) + \int_{\pi-0}^{0+0} (2\pi-\phi)\, d\tau_{pq}(2\pi-\phi) = 0,$$

inasmuch as the values of $\omega_{pq}(\phi)$ at its inner discontinuity points play no rôle in the Stieltjes integration in (68) and (69). Formulas (64) and (65) obviously follow from (66) and (67) by virtue of (68) and (69), respectively. The reduced Hermitian matrix H in the exponential representation of Theorem I may always be calculated from the spectral matrix of the given unitary matrix U by means of (56). For the treatment of a real unitary matrix $U - \bar{U} = 0$ we express the Stieltjes integration in (56) as follows:

$$(70)\qquad 2\pi H = \| \int_{0+0}^{\pi-0} \phi d\tau_{pq}(\phi) \| + \pi \| \underset{\pi}{\Delta}\, \tau_{pq}(\phi) \| + \| \int_{\pi+0}^{2\pi-0} \phi d\tau_{pq}(\phi) \|,$$

inasmuch as [cf. (43′) on p. 598]

$$\int_{0}^{0+0} \phi d\tau_{pq}(\phi) = \int_{2\pi-0}^{2\pi} \phi d\tau_{pq}(\phi) = 0.$$

By means of (65) we may rewrite (70) in the form

$$2\pi H = \| \int_{0+0}^{\pi-0} \phi d[\tau_{pq}(\phi) - \bar{\tau}_{pq}(\phi)] \| + \pi \| \underset{\pi}{\Delta} \tau_{pq}(\phi) \| + 2\pi \| \int_{0+0}^{\pi-0} d\bar{\tau}_{pq}(\phi) \|,$$

which in turn may be replaced by

$$(71)\quad 2\pi\mathrm{H} = \left\| \int_{0+0}^{\pi-0} \phi d[\tau_{pq}(\phi) - \bar{\tau}_{pq}(\phi)] \right\| + \pi \left\| \int_{0+0}^{\pi-0} d[\bar{\tau}_{pq}(\phi) - \tau_{pq}(\phi)] \right\|$$
$$+ \pi \left\| \int_{0+0}^{\pi-0} d\tau_{pq}(\phi) \right\| + \pi \left\| \Delta_{\pi} \tau_{pq}(\phi) \right\| + \pi \left\| \int_{0+0}^{\pi-0} d\bar{\tau}_{pq}(\phi) \right\|.$$

We introduce a bounded real skew-symmetric matrix S by means of

$$(72)\quad 2\pi\mathrm{S} = -i \left\| \int_{0+0}^{\pi-0} \phi d[\tau_{pq}(\phi) - \bar{\tau}_{pq}(\phi)] \right\| - i\pi \left\| \int_{0+0}^{\pi-0} d[\bar{\tau}_{pq}(\phi) - \tau_{pq}(\phi)] \right.$$
$$= -i \left\| \int_{0+0}^{\pi-0} (\phi - \pi) d[\tau_{pq}(\phi) - \bar{\tau}_{pq}(\phi)] \right\|,$$

and employ it, together with (64), to abbreviate (71) to

$$(73)\qquad 2\mathrm{H} = 2i\mathrm{S} + \left\| \int_{0+0}^{\pi-0} d\tau_{pq}(\phi) \right\| + \left\| \Delta_{\pi} \tau_{pq}(\phi) \right\| + \left\| \int_{\pi+0}^{2\pi-0} d\tau_{pq}(\phi) \right\|.$$

Now from (38′) [cf. p. 598] for $n = 0$ we have

$$2\pi\mathrm{E} = \left\| \int_{0}^{2\pi} d\tau_{pq}(\phi) \right\| = \left\| \Delta_{0+0} \tau_{pq}(\phi) \right\|$$
$$+ \left\| \int_{0+0}^{\pi-0} d\tau_{pq}(\phi) \right\| + \left\| \Delta_{\pi} \tau_{pq}(\phi) \right\| + \left\| \int_{\pi+0}^{2\pi-0} d\tau_{pq}(\phi) \right\|,$$

inasmuch as from (43′)

$$\int_{2\pi-0}^{2\pi} d\tau_{pq}(\phi) = 0,$$

so that (73) may be further shortened to

$$(74)\quad \mathrm{H} = i\mathrm{S} + \pi\mathrm{E} - (1/2) \left\| \Delta_{0+0} \tau_{pq}(\phi) \right\| = i\mathrm{S} + \pi[\mathrm{E} - (1/2\pi) \left\| \Delta_{0+0} \tau_{pq}(\phi) \right\|]$$

in which S is the bounded real skew-symmetric matrix introduced in (72) and $\| \Delta_{0+0} \tau_{pq}(\phi) \|$, since $\| \tau_{pq}(\phi) \|$ is a bounded Hermitian matrix, is, from II on p. 599, a bounded real symmetric matrix. We introduce a bounded real [cf. (61)] symmetric matrix B by means of

$$\mathrm{B} = \mathrm{E} - (1/2\pi) \left\| \Delta_{0+0} \tau_{pq}(\phi) \right\|,$$

and which, from (59), may obviously be written in the form

$$(75)\qquad \mathrm{B} = \mathrm{E} - \left\| \Delta_{0+0} \sigma_{pq}(\phi) \right\|,$$

so that, inasmuch as $\| \sigma_{pq}(\phi) \|$ is the spectral matrix of the reduced Hermitian matrix H we have †

† See "Hi", pp. 141-145. Cf. also "S", p. 160 and p. 174.

$$B^2 = E - 2 \| \underset{0+0}{\Delta} \sigma_{pq}(\phi) \| + \| \underset{0+0}{\Delta} \sigma_{pq}(\phi) \|^2 = E - \| \underset{0+0}{\Delta} \sigma_{pq}(\phi) \| = B,$$

.e. B is a real symmetric idempotent matrix so that the necessity of (63a) s demonstrated.

In order to demonstrate the necessity of (63b) we write (72) in the form

$$(76) \qquad 2\pi S = \lim_{n\to+\infty} 2\pi S_n,$$

where

$$2\pi S_n = -i \left\| \sum_{\nu=1}^{n-2} \left(\frac{\nu}{n}\pi - \pi\right) \left\{ \left[\tau_{pq}\left(\frac{\nu+1}{n}\pi\right) - \tau_{pq}\left(\frac{\nu}{n}\pi\right)\right] - \left[\bar{\tau}_{pq}\left(\frac{\nu+1}{n}\pi\right) - \bar{\tau}_{pq}\left(\frac{\nu}{n}\pi\right)\right] \right\} \right\|.$$

, according to (59),

$$(77) \quad S_n = -i \left\| \sum_{\nu=1}^{n-2} \left(\frac{\nu}{n}\pi - \pi\right) \left\{ \left[\sigma_{pq}\left(\frac{\nu+1}{n}\pi\right) - \sigma_{pq}\left(\frac{\nu}{n}\pi\right)\right] - \left[\bar{\sigma}_{pq}\left(\frac{\nu+1}{n}\pi\right) - \bar{\sigma}_{pq}\left(\frac{\nu}{n}\pi\right)\right] \right\} \right\|.$$

We now apply the well-known theorem of Hilbert † that

$$\| \sigma_{pq}(\phi_1) \| \; \| \sigma_{pq}(\phi_2) \| = \| \sigma_{pq}(\phi_2) \| \; \| \sigma_{pq}(\phi_1) \|,$$

o obtain

$$(78) \qquad \| \underset{0+0}{\Delta} \sigma_{pq}(\phi) \| \; \| \sigma_{pq}(\nu\pi/n) \| = \| \sigma_{pq}(\nu\pi/n) \| \; \| \underset{0+0}{\Delta} \sigma_{pq}(\phi) \|; \quad (\nu = 1, 2, \cdots, n-1); \; (n = 2, 3, \cdots),$$

and hence also

$$(79) \qquad \| \underset{0+0}{\Delta} \sigma_{pq}(\phi) \| \; \| \bar{\sigma}_{pq}(\nu\pi/n) \| = \| \bar{\sigma}_{pq}(\nu\pi/n) \| \; \| \underset{0+0}{\Delta} \sigma_{pq}(\phi) \|; \quad (\nu = 1, 2, \cdots, n-1); \; (n = 2, 3, \cdots),$$

nasmuch as, from II on p. 599, the matrix $\| \underset{0+0}{\Delta} \sigma_{pq}(\phi) \|$ is a real matrix.

From (75), (77), (78) and (79) we accordingly have

$$(E - B)S_n = S_n(E - B),$$

from which and (76) then follows, for $n \to +\infty$, the necessity of (63b).

We now demonstrate the converse part of Theorem I, i. e. we assume that $H = \pi B + iS$ is a reduced Hermitian matrix for which the matrices

† See "Hi", pp. 141-145.

B and S fulfill the conditions (63a) and (63b) and demonstrate that U is a real matrix. We precede the proof by the derivation of the identity

$$e^{i\theta I} = E + (e^{i\theta} - 1)I, \tag{80}$$

where θ is an arbitrary real or complex number and I any bounded idempotent matrix (cf. p. 585). Since I is assumed to be bounded we have the convergent development (cf. p. 588)

$$e^{i\theta I} = \sum_{\nu=0}^{+\infty} \frac{(i\theta I)^{\nu}}{\nu!} = \sum_{\nu=0}^{+\infty} \frac{(i\theta I)^{2\nu}}{(2\nu)!} + \sum_{\nu=0}^{+\infty} \frac{(i\theta I)^{2\nu+1}}{(2\nu+1)!},$$

which may, since I is idempotent, obviously be written in the form

$$e^{i\theta I} = E - I + \left[\sum_{\nu=0}^{+\infty} \frac{(i\theta)^{2\nu}}{(2\nu)!} + \sum_{\nu=0}^{+\infty} \frac{(i\theta)^{2\nu+1}}{(2\nu+1)!}\right] I = E - I + \left[\sum_{\nu=0}^{+\infty} \frac{(i\theta)^{\nu}}{\nu!}\right] I$$

from which (80) follows at once. In an analogous manner it is possible to obtain the identity

$$e^{i\theta Q} = E \cos\theta + iQ \sin\theta,$$

valid for arbitrary real or complex θ and for an arbitrary square root Q of the unit matrix E (cf. p. 585). From the commutability condition (63b) we obtain (62′) from which if we apply (80) for $\theta = \pi$ and $I = B$ we obtain

$$e^{iH} = (E - 2B) \cdot e^{-S} = e^{-S} \cdot (E - 2B),$$

which is, because of (63a), certainly a real matrix. q. e. d.

III. A Connection Between Reduced Skew-Symmetric and Reduced Hermitian Matrices.

For convenience we adopt the notation

$$M \sim N \tag{82}$$

for the orthogonal equivalence of two matrices, i. e. for the existence of an orthogonal matrix O for which $M = ONO^{-1}$. We introduce the definition

A real skew-symmetric matrix S *will be said to be reduced if the spectrum of the Hermitian matrix* iS *is contained in the closed interval* $[-\pi, +\pi]$ *and if the end-points* $-\pi$ *and* $+\pi$ *are not contained in the point spectrum of* iS.

The property of a real skew-symmetric matrix to be reduced is an obvious orthogonal invariant.

We now show

LEMMA I. *Let* B *be a real symmetric idempotent matrix and let* S *be a real bounded skew-symmetric matrix commutable with* B, *i. e.*

$$\text{(83)} \qquad BS = SB,$$

then the necessary and sufficient conditions in order that the Hermitian matrix $\pi B + iS$ *be reduced* (*cf. p.* 588) *are*

$$\text{(84a)} \qquad (E - B)S = S(E - B) = \| 0 \|,$$

$$\text{(84b)} \qquad S \textit{ be reduced.}$$

Since B is a real symmetric idempotent matrix we have, according to a theorem of Hilbert,† in the notation (82)

$$\text{(85)} \quad B \sim \| \kappa_{pq} \| \text{ where } \kappa_{pq} = 0 \text{ for } p \neq q \text{ and } \kappa_{pp} = 0 \text{ or } \kappa_{pp} = +1.$$

We designate by α' and β' the number of times 0 and $+1$ respectively occur in the sequence $\{\kappa_{pp}\}$ so that necessarily $\alpha' \geqq 0$, $\beta' \geqq 0$ and $\alpha' + \beta' = +\infty$ and distinguish the following four possibilities for the matrix $\| \kappa_{pq} \|$ in (85):

$$\text{(86)} \quad \begin{array}{rlll} \text{I:} & \alpha' = +\infty, & \beta' = 2m & \text{for } 0 \leqq m < +\infty, \\ \text{II:} & \alpha' = +\infty, & \beta' = 2m+1 & \text{for } 0 \leqq m < +\infty, \\ \text{III:} & \alpha' = m, & \beta' = +\infty & \text{for } 0 \leqq m < +\infty, \\ \text{IV:} & \alpha' = +\infty, & \beta' = +\infty, & \end{array}$$

and corresponding to these four possibilities for the matrix $\| \kappa_{pq} \|$ in (85) we have the following four canonical forms for the matrix B:

$$\text{(87)} \quad \begin{array}{rl} \text{I:} & B \sim B^{(1)} = \| b^{(1)}_{pq} \|, \text{ where } b^{(1)}_{pq} = 0 \text{ for } p \neq q, \\ & \quad \text{and } b^{(1)}_{pp} = 1 \text{ for } p = 1, 2, \cdots, 2m < +\infty, \\ & \quad b^{(1)}_{pp} = 0 \text{ for } p > 2m, \\ \text{II:} & B \sim B^{(2)} = \| b^{(2)}_{pq} \|, \text{ where } b^{(2)}_{pq} = 0 \text{ for } p \neq q, \\ & \quad \text{and } b^{(2)}_{pp} = 1 \text{ for } p = 1, 2, \cdots, 2m+1 < +\infty, \\ & \quad b^{(2)}_{pp} = 0 \text{ for } p > 2m+1, \\ \text{III:} & B \sim B^{(3)} = \| b^{(3)}_{pq} \|, \text{ where } b^{(3)}_{pq} = 0 \text{ for } p \neq q, \\ & \quad \text{and } b^{(3)}_{pp} = 0 \text{ for } p = 1, 2, \cdots, m < +\infty, \\ & \quad b^{(3)}_{pp} = 1 \text{ for } p > m, \\ \text{IV:} & B \sim B^{(4)} = \| b^{(4)}_{pq} \|, \text{ where } b^{(4)}_{pq} = 0 \text{ for } p \neq q, \\ & \quad \text{and } b^{(4)}_{pp} = b^{(4)}_{p+1\,p+1} = 1 \text{ for } p = 1, 5, 9, \cdots, \\ & \quad b^{(4)}_{pp} = b^{(4)}_{p+1\,p+1} = 0 \text{ for } p = 3, 7, 11, \cdots. \end{array}$$

† See "Hi", p. 145.

Since the commutability property (83) assumed for the matrices B and S in Lemma I is invariant under simultaneous transformations of B and S we must have, simultaneously with the matrix transformation of the matrix B of Lemma I into each of the above canonical forms, the reduction of the skew-symmetric matrix S of Lemma I to one of the following types:

$$
(88)\quad
\begin{aligned}
&\mathrm{I}:\quad \mathrm{S} \sim \mathrm{S}^{(1)} = \left\| \begin{array}{c|c} S_{2m} & \\ \hline & \mathrm{S}_{\infty-2m} \end{array} \right\|, \\
&\qquad S_{2m} = \bar{S}_{2m} = -S'_{2m}, \quad \mathrm{S}_{\infty-2m} = \bar{\mathrm{S}}_{\infty-2m} = -\mathrm{S}'_{\infty-2m}, \\
&\mathrm{II}:\quad \mathrm{S} \sim \mathrm{S}^{(2)} = \left\| \begin{array}{c|c} S_{2m+1} & \\ \hline & \mathrm{S}_{\infty-(2m+1)} \end{array} \right\|, \\
&\qquad S_{2m+1} = \bar{S}_{2m+1} = -S'_{2m+1}, \quad \mathrm{S}_{\infty-(2m+1)} = \bar{\mathrm{S}}_{\infty-(2m+1)} = -\mathrm{S}'_{\infty-(2m+1)}, \\
&\mathrm{III}:\quad \mathrm{S} \sim \mathrm{S}^{(3)} = \left\| \begin{array}{c|c} S_{m} & \\ \hline & \mathrm{S}_{\infty-m} \end{array} \right\|, \\
&\qquad S_{m} = \bar{S}_{m} = -S'_{m}, \quad \mathrm{S}_{\infty-m} = \bar{\mathrm{S}}_{\infty-m} = -\mathrm{S}'_{\infty-m}, \\
&\mathrm{IV}:\quad \mathrm{S} \sim \mathrm{S}^{(4)} = \left\| \begin{array}{ccccccc} S_{11} & Z & S_{13} & Z & \cdot & \cdot & \cdot \\ Z & S_{22} & Z & S_{23} & \cdot & \cdot & \cdot \\ S_{31} & Z & S_{33} & Z & \cdot & \cdot & \cdot \\ Z & S_{42} & Z & S_{44} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & \cdot & & & \end{array} \right\|, \\
&\qquad S_{pq} = \bar{S}_{pq} = -S'_{qp} = \left\| \begin{array}{cc} \alpha_{pq} & \beta_{pq} \\ \gamma_{pq} & \delta_{pq} \end{array} \right\|, \quad Z = \left\| \begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array} \right\|.
\end{aligned}
$$

in which the capital letters represent matrices and the vacant spaces of the remaining two "quadrants" of I, II and III are occupied entirely by zeros. For example, the matrix $\mathrm{S}^{(1)} = \| s^{(1)}_{pq} \|$ of (88) is a bounded real skew-symmetric matrix for which

$$
(89)\qquad s^{(1)}_{pq} = 0, \quad \text{for} \quad p = 1, 2, \cdots, 2m,\ 0 \leqq 2m < q; \quad \text{and for} \quad 0 \leqq 2m < p,\ q = 1, 2, \cdots, 2m.
$$

We now prove the necessity of the conditions (84a), (84b) in Lemma I. As mentioned in connection with Theorem I the property of a Hermitian matrix to be reduced is an orthogonal invariant. Consequently the Hermitian matrices

$$
(90)\qquad \pi\mathrm{B}^{(j)} + i\mathrm{S}^{(j)}; \qquad (j = 1, 2, 3, 4),
$$

constructed from (87) and (88), are all reduced. In order to show the necessity of the conditions (84a), (84b) of Lemma I we consider separately each of the four possible canonical forms for the matrix B of Lemma I as given in (87).

1. *Proof of the necessity of* (84a) *for matrices* B *of Class I.* From I of (88) we are enabled to write

$$(91)\qquad S^{(1)} = S_a^{(1)} + S_b^{(1)},$$

where, in the schematic notations of (88), (89), we have placed

$$(92)\qquad S_a^{(1)} = \left\| \begin{array}{c:c} S_{2m} & \\ \hdashline & \end{array} \right\|, \qquad S_b^{(1)} = \left\| \begin{array}{c:c} & \\ \hdashline & S_{\infty-2m} \end{array} \right\|.$$

so that the matrix pair $(S_a^{(1)}, S_b^{(1)})$ is separated (cf. p. 582). Since the Hermitian matrix $\pi B^{(1)} + iS^{(1)}$ is reduced [cf. (90) and adjoining] the domain $\Delta(\pi B^{(1)} + iS^{(1)})$ lies in the closed interval $[0, 2\pi]$ (cf. p. 582), i. e. from (91)

$$(93)\qquad \begin{aligned} 0 \leqq \Phi[(\pi B^{(1)} + iS^{(1)}); x] &= \pi\Phi[B^{(1)}; x] \\ &+ \Phi[iS_a^{(1)}; x] + \Phi[iS_b^{(1)}; x] \leqq 2\pi \quad \text{for} \quad |x| = 1. \end{aligned}$$

Since (93) holds for an arbitrary unit vector x we may choose in (93)

$$(94)\qquad x_1 = x_2 = \cdots = x_{2m} = 0, \qquad \sum_{\nu=2m+1}^{+\infty} |x_\nu|^2 = 1,$$

so that, from (92) and the definition of $B^{(1)}$ in (87), we derive from (93)

$$(95)\qquad 0 \leqq \Phi[iS_b^{(1)}; x] \leqq 2\pi,$$

first, for only those unit vectors x of (94), and then, because of the peculiar definition of $S_b^{(1)}$ given in (92), for an arbitrary unit vector x. Since the domain of any bounded Hermitian matrix of the type iS, where S is a real skew-symmetric matrix, is necessarily symmetrical with respect to the origin (cf. p. 582) one must perforce replace (95) by

$$(96)\qquad \Phi[iS_b^{(1)}; x] \equiv 0 \quad \text{for} \quad |x| = 1,$$

and therefore from the uniqueness property of compound forms (cf. p. 580)

$$S_b^{(1)} = \| 0 \|,$$

or from (91)

$$(97)\qquad S^{(1)} = S_a^{(1)}.$$

From (87), (97), and (92) it is clear that the matrix pair $(E - B^{(1)}, S^{(1)})$

is separated and consequently *a fortiori* perpendicular. Since the perpendicularity of a matrix pair is invariant with respect to simultaneous matrix transformations of the two matrices in the pair (cf. p. 583) the matrix pair $(E - B, S)$ of Lemma I is itself perpendicular and the necessity of (84a) is accordingly demonstrated, under the temporary restriction that B be of Class I.

2. *Proof of the necessity of* (84b) *for matrices* B *of Class I.* We choose instead of (94) those unit vectors x for which

$$(98)\qquad \sum_{\nu=1}^{2m} | x_\nu |^2 = 1, \qquad x_{2m+1} = x_{2m+2} = \cdots = 0,$$

so that, from (92) and the definition of $B^{(1)}$ in (87) we derive from (93)

$$(99)\qquad 0 \leqq \pi + \Phi\,[iS_a{}^{(1)}; x] \leqq 2\pi, \quad \text{i.e.} \quad -\pi \leqq \Phi\,[iS_a{}^{(1)}; x] \leqq +\pi,$$

first, for only those unit vectors of (98), and then because of the peculiar definition of $S_a{}^{(1)}$ given in (92), for an arbitrary unit vector x. Since, as has been mentioned in the introduction (cf. p. 581), the spectrum of any bounded matrix is contained in the domain of the matrix it is clear from (99) that the spectrum of $iS_a{}^{(1)}$ is contained in the closed interval $[-\pi, +\pi]$. Furthermore the matrix $E - B^{(1)}$ is a diagonal matrix the elements of which are 0 and $+1$ and forms with $S^{(1)}$ the separated matrix pair $(E - B^{(1)}, S^{(1)})$ [cf. (87); (97), (92)] so that from p. 584 and the fact that the Hermitian matrix $\pi B^{(1)} + iS^{(1)}$ is reduced [cf. (90) and adjoining] the points $-\pi$ and $+\pi$ cannot be in the point spectrum of $iS^{(1)}$, i. e. $S^{(1)}$ is a reduced skew-symmetric matrix. Since the property of a skew-symmetric matrix to be reduced is an orthogonal invariant the skew-symmetric matrix S of Lemma I is itself reduced and the necessity of (84b) is, under the temporary restriction that B be of type I of (87), accordingly demonstrated.

3. *Proof of the necessity of* (84a) *and* (84b) *for the matrices* B *of Classes II, III, and IV.* That (84a) and (84b) are necessary if the matrix B in Lemma I is of type II of (87) follows at once on replacing $2m$ by $2m + 1$ throughout the preceding demonstrations. In order to treat type III of (87) we may write, analogously to (91) and (92), from (88)

$$(100)\qquad S^{(3)} = S_a{}^{(3)} + S_b{}^{(3)},$$

where

$$(101)\qquad S_a{}^{(3)} = \left\| \begin{array}{c|c} S_m & \\ \hline & \end{array} \right\|, \qquad S_b{}^{(3)} = \left\| \begin{array}{c|c} & \\ \hline & S_{\infty - m} \end{array} \right\|.$$

so that the matrix pair $(S_a^{(3)}, S_b^{(3)})$ is separated. Since the matrix $\pi B^{(3)} + iS^{(3)}$ is reduced [cf. (90) and adjoining] we accordingly have (cf. p. 582) from (100)

$$(102)\quad 0 \leqq \Phi\,[(\pi B^{(3)} + iS^{(3)})\,;x] \\ = \pi\Phi\,[B^{(3)};x] + \Phi\,[iS_a^{(3)};x] + \Phi\,[iS_b^{(3)};x] \leqq 2\pi \text{ for } |\,x\,| = 1,$$

and hence

$$(103)\quad 0 \leqq \pi + \Phi[iS_b^{(3)};x] \leqq 2\pi \text{ for } x_1 = x_2 = \cdots x_m = 0,\ \sum_{\nu=m+1}^{+\infty} |\,x_\nu\,|^2 = 1,$$

$$(104)\quad 0 \leqq \Phi[iS_a^{(3)};x] \leqq 2\pi \text{ for } \sum_{\nu=1}^{m} |\,x_\nu\,|^2 = 1, \quad x_{m+1} = x_{m+2}\cdots = 0,$$

so that one may deduce [cf. (95), (99) and adjoining] from (103) and (104) that the matrix pair $(E - B^{(3)}, S^{(3)})$ is separated and that the skew-symmetric matrix $S^{(3)}$ is reduced and then, as a consequence, that the matrix pair $(E - B, S)$ of Lemma I is perpendicular and that the skew-symmetric matrix S is reduced. In order to treat the remaining case in which the matrix B of Lemma I is of the type IV of (87) we write

$$(105)\qquad S^{(4)} = S_a^{(4)} + S_b^{(4)},$$

where, in the notations employed in (88), we understand

$$(106)\quad S_a^{(4)} = \left\| \begin{matrix} \boldsymbol{S}_{11} & \boldsymbol{Z} & \boldsymbol{S}_{13} & \boldsymbol{Z} & \cdots \\ \boldsymbol{Z} & \boldsymbol{Z} & \boldsymbol{Z} & \boldsymbol{Z} & \cdots \\ \boldsymbol{S}_{31} & \boldsymbol{Z} & \boldsymbol{S}_{33} & \boldsymbol{Z} & \cdots \\ \boldsymbol{Z} & \boldsymbol{Z} & \boldsymbol{Z} & \boldsymbol{Z} & \cdots \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \end{matrix} \right\|, \qquad S_b^{(4)} = \left\| \begin{matrix} \boldsymbol{Z} & \boldsymbol{Z} & \boldsymbol{Z} & \boldsymbol{Z} & \cdots \\ \boldsymbol{Z} & \boldsymbol{S}_{22} & \boldsymbol{Z} & \boldsymbol{S}_{24} & \cdots \\ \boldsymbol{Z} & \boldsymbol{Z} & \boldsymbol{Z} & \boldsymbol{Z} & \cdots \\ \boldsymbol{Z} & \boldsymbol{S}_{42} & \boldsymbol{Z} & \boldsymbol{S}_{44} & \cdots \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \end{matrix} \right\|,$$

so that the matrix pair $(S_a^{(4)}, S_b^{(4)})$ is separated. Since the Hermitian matrix $\pi B^{(4)} + iS^{(4)}$ is reduced [cf. (90) and adjoining] we accordingly have (cf. p. 582) from (105)

$$(107)\quad 0 \leqq \Phi\,[(\pi B^{(4)} + iS^{(4)};x] = \pi\Phi\,[B^{(4)};x] \\ + \Phi\,[iS_a^{(4)};x] + \Phi\,[iS_b^{(4)};x] \leqq 2\pi \quad \text{for} \quad |\,x\,| = 1,$$

and since x is an arbitrary unit vector in (107) we may put

$$(108)\qquad x_1 = x_2 = x_5 = x_6 = x_9 = x_{10} = \cdots = 0; \\ |\,x_3\,|^2 + |\,x_4\,|^2 + |\,x_7\,|^2 + |\,x_8\,|^2 + \cdots = 1,$$

so that from (106) and the definition of $B^{(4)}$ in (87) we may derive from (107)

$$0 \leqq \Phi\,[iS_b^{(4)}; x] \leqq 2\pi, \tag{109}$$

first, for only those unit vectors x in (108), and then, because of the peculiar definition of $S_b^{(4)}$ given in (106), for an arbitrary unit vector x. A treatment of inequality (109) analogous to that given for (95) then yields

$$S_b^{(4)} = \| 0 \|,$$

or from (105)

$$S^{(4)} = S_a^{(4)}. \tag{110}$$

From (87), (110), and (106) it is clear that the matrix pair $(E - B^{(4)}, S^{(4)})$ is separated and consequently perpendicular. From this point on the proof of the necessity of (84a) for those matrices B of Lemma I of type IV in (87) is identical with the proof given in the treatment of type I and is accordingly not repeated.

In order to show the necessity of (84b) if the matrix B in Lemma I is of type IV in (87) we choose, instead of (108), those unit vectors x for which

$$|x_1|^2 + |x_2|^2 + |x_5|^2 + |x_6|^2 + |x_9|^2 + |x_{10}|^2 = \cdots = 1, \quad x_3 = x_4 = x_7 = x_8 = \cdots = 0 \tag{111}$$

so that now, from (106) and the definition of $B^{(4)}$ in (87), the inequality (107) is replaced by

$$0 \leqq \pi + \Phi\,[iS_a^{(4)}; x] \leqq 2\pi, \quad \text{i.e.} \quad -\pi \leqq \Phi\,[iS_a^{(4)}; x] \leqq +\pi, \tag{112}$$

first, only for those unit vectors x in (111), and then, because of the peculiar definition of $S_a^{(4)}$ given in (106), for an arbitrary unit vector x. By employing the same argument used in the discussion of (99) we learn that the spectrum of $iS_a^{(4)}$ is contained in the closed interval $[-\pi, +\pi]$ and that neither of the end-points $-\pi$ and $+\pi$ are contained in the point spectrum of $iS_a^{(4)}$ (cf. p. 610). Finally it follows, just as in the treatment on p. 610, that the matrix S in Lemma I is a reduced skew-symmetric matrix.

4. *Sufficiency of the conditions* (84a) *and* (84b) *in Lemma I.* We now turn to the proof of the sufficiency of (84a) and (84b) in Lemma I. Since B is now by hypothesis a real symmetric idempotent matrix and S a real skew-symmetric matrix commutable with B [the commutability condition (83) is a necessary but not a sufficient condition for the perpendicularity of the matrix pair $(E - B, S)$], we may by *simultaneous* orthogonal matrix transformations obtain the canonical forms in (87) and (88) for the matrices B and S, respectively. As has been mentioned before, the assumptions in

(84a) and (84b) for the matrices B and S are orthogonal invariants and since the property of a Hermitian matrix to be reduced is likewise an orthogonal invariant it will be sufficient to prove the sufficiency for the Hermitian matrices in (90). We consider separately the four possibilities in (87) and begin with the treatment of type I. From the perpendicularity of the matrix pair $(E - B^{(1)}, S^{(1)})$ there follows from (87) and (88) that $S_b^{(1)} = \| 0 \|$ in (92), i. e. from (91), the matrix equation (97). Since $S^{(1)}$ is a reduced skew-symmetric matrix it follows from (97), (92) and (87) that $\pi B^{(1)} + iS^{(1)}$ is a reduced Hermitian matrix (cf. p. 584); from which the sufficiency of (84a) and (84b) in Lemma I follows [if B is of type I in (87)]. The treatment of the types II and III of (87) obviously proceeds entirely along the same lines and accordingly is not repeated here. The treatment of type IV of (87), while proceeding in the same general manner, is however more complicated and consequently is given in some detail. From the perpendicularity of the matrix pair $(E - B^{(4)}, S^{(4)})$ there follows from (87) and (88) that $S_b^{(4)} = \| 0 \|$ in (106), i. e. from (105), the matrix equation (110). Since $S^{(4)}$ is a reduced skew-symmetric matrix it follows from (110), (106), and (87) that first, the spectrum of $\pi B^{(4)} + iS^{(4)}$ is contained in the closed interval $[0, 2\pi]$ and second, from p. 584, that the end-point 2π is not contained in the point spectrum of $\pi B^{(4)} + iS^{(4)}$; from which the sufficiency of (84a) in Lemma I for the matrices B of the type IV in (87) follows. q. e. d.

III. The Parametric Representation of the Manifold of the Orthogonal Matrices.

We first employ Theorem I and Lemma I for the demonstration of the following two theorems:

Theorem II. *Every orthogonal matrix* O *may be represented in the form*

$$(113) \qquad O = e^{i(\pi B + iS)} = e^{i\pi B} \cdot e^{-S} = e^{-S} \cdot e^{i\pi B},$$

in which B *is a real symmetric idempotent matrix and* S *a real bounded skew-symmetric matrix commutable with* B, *i. e.*

$$(114a) \qquad B = \bar{B} = B', \; B^2 = B; \quad S = \bar{S} = -S',$$

$$(114b) \qquad BS = SB.$$

Furthermore this representation is possible and unique if the matrix pair

$(E-B, S)$ *is perpendicular and if the skew-symmetric matrix* S *is reduced, i. e. if*

$$(115a)\qquad (E-B)S = S(E-B) = \| 0 \|,$$

$$(115b)\qquad S \textit{ is reduced.}$$

THEOREM III. *Every orthogonal matrix* O *may be represented in the form*

$$(116)\qquad O = 2C - e^{-S},$$

in which C *is a real symmetric idempotent matrix and* S *a bounded real skew-symmetric matrix such that the matrix pair* (C, S) *is perpendicular, i. e.*

$$(117a)\qquad C = \bar{C} = C', C^2 = C; \quad S = \bar{S} = -S',$$

$$(117b)\qquad CS = SC = \| 0 \|,$$

and this representation is unique if the skew-symmetric matrix S *in* (116) *is reduced. The matrix* C *is in fact the matrix* $E-B$ *of Theorem II.*

Theorem II is a trivial consequence of Theorem I and Lemma I in that for its demonstration it is only necessary to replace the condition in Theorem I that the Hermitian matrix $\pi B + iS$ be reduced by the equivalent conditions (84a) and (84b) of Lemma I. That the representation (113), (114a), (114b) always yields an orthogonal matrix, i. e. a real unitary matrix even without the conditions (115a), (115b) is trivial inasmuch as (cf. p. 606)

$$O = e^{i(\pi B + iS)} = e^{i\pi B} \cdot e^{-S} = (E-2B)e^{-S} = \bar{O},$$

from which it is clear that the prescribing of the conditions (115a) and (115b) is responsible for the uniqueness of the representation (113).

Theorem III is now demonstrated as a consequence of Theorem II and the identity (19) on p. 588 valid for every perpendicular pair (M, N) of bounded matrices. In order to show this we write the representation (113) as follows:

$$(118)\qquad \begin{aligned} O = \exp\,[i(\pi B + iS)] &= \exp\,[i(\pi E - \pi(E-B) + iS)] \\ &= -\exp i\,[-\pi(E-B) + iS]. \end{aligned}$$

If we now choose the matrices B and S to fulfill the conditions (114a), (114b), (115a), (115b) the representation (113), (118) provides a unique representation of any given orthogonal matrix O and which may, from (19), be written in the form

$$O = E - e^{-i\pi(E-B)} - e^{-S},$$

which in turn, from (80) [$\theta = -\pi$ and $I = E - B$], may be written as

$$O = 2(E - B) - e^{-S}, \tag{119}$$

inasmuch as for $B^2 = B$, as is well known,

$$(E - B)^2 = E^2 - 2B + B^2 = E - B.$$

The matrix $E - B$ in (119) may accordingly be identified as the matrix C in Theorem III. Finally it is clear that the representation (116), (117a), (117b) always yields an orthogonal matrix, even though S is not a reduced skew-symmetric matrix, for

$$\begin{aligned} OO' = O'O &= [2(E - B) - e^{-S}]\,[2(E - B) - e^{S}] \\ &= 4(E - B)^2 - 2(E - B)e^{S} - 2e^{-S}(E - B) + E \\ &= 4(E - B) - 2(E - B) - 2(E - B) + E = E. \end{aligned}$$

It is clear from the above considerations that the only effect on the representation (116) of prescribing that the skew-symmetric matrix S be reduced is to insure the uniqueness of the representation.

We now show

LEMMA II. *The formula* $Q = E - 2I$ *yields a one-to-one correspondence between the idempotent matrices* I *and the matrices* Q *which are square roots of* E. *This one-to-one correspondence has the obvious property that the matrices* I *and* Q *in it are simultaneously real and simultaneously symmetric.*

Let I be a given arbitrary (not necessarily Hermitian) idempotent matrix, i. e. $I^2 = I$. The matrix $Q = E - 2I$ is then uniquely determined and satisfies the matrix equation

$$Q^2 = (E - 2I)^2 = E^2 - 4I + 4I^2 = E,$$

so that Q is a square root of E.

Let Q be an arbitrary (not necessarily Hermitian) square root of E, i. e. $Q^2 = E$. The matrix $I = \frac{1}{2}(E - Q)$ is then uniquely determined and satisfies the matrix equation

$$I^2 = (1/4)(E - Q)^2 = (1/4)(E^2 - 2Q + Q^2) = (1/2)(E - Q) = I,$$

so that I is an idempotent matrix.

That the correspondence $Q = E - 2I$ is one to one is trivial.

As a consequence of Theorem I and Lemma II we show

THEOREM IV. *Every orthogonal matrix* O *may be represented in the form*

(120) $$O = Qe^{-S} = e^{-S}Q,$$

in which Q *is a real symmetric square root of* E (*and therefore itself an orthogonal matrix inasmuch as* $QQ' = Q'Q = Q^2 = E$), S *a bounded real skew-symmetric matrix, commutable with* Q, *i. e.*

(121a) $$Q = \bar{Q} = Q', \ Q^2 = E; \quad S = \bar{S} = -S',$$

(121b) $$QS = SQ,$$

and this correspondence between the orthogonal matrices O *and the matrix pairs* (Q, S) *for which the Hermitian matrices* $(\pi/2)(E - Q) + iS$ *are reduced is one-to-one.*

The matrix B in the representation (62) is, from (63a), a real symmetric idempotent matrix. According to the above lemma there exists a uniquely determined real symmetric square root Q of E for which (62) may be written in the form

(122) $$O = \exp i[(\pi/2)(E - Q) + iS] = \exp [i(\pi/2)(E - Q)] \cdot \exp [-S].$$

The matrix $B = \frac{1}{2}(E - Q)$ is idempotent so that (122) may, from (80) [$\theta = \pi$ and $I = (1/2)(E - Q)$], be written in the desired form (120). Finally from Lemma II and Theorem I the correspondence (120), (122) between the orthogonal matrices O and the matrix pairs (Q, S) of (121a), (121b) is one-to-one provided that the Hermitian matrix $\pi B + iS$, i. e. $(\pi/2)(E - Q) + iS$ is reduced. q. e. d.

V. The Four Classes of Orthogonal Matrices.

According to Lemma II if Q is a real symmetric square root of E there exists a uniquely determined real symmetric idempotent matrix B for which

(123) $$Q = E - 2B$$

holds. Corresponding to the four canonical forms for the matrices B, under orthogonal matrix transformations, given in (87) we accordingly obtain the following four canonical forms for the matrices Q which are square roots of E:

$$\begin{aligned}
&\mathrm{I}: \ Q \sim Q^{(1)} = \| q^{(1)}{}_{jk} \|, \ q^{(1)}{}_{jk} = 0 \ j \neq k, \ q^{(1)}{}_{jj} = -1 \\
&\qquad \text{for } j = 1, 2, \cdots, 2m < +\infty, \ q^{(1)}{}_{jj} = +1 \text{ for } j > 2m, \\
&\mathrm{II}: \ Q \sim Q^{(2)} = \| q^{(2)}{}_{jk} \|, \ q^{(2)}{}_{jk} = 0 \ j \neq k, \ q^{(2)}{}_{jj} = -1 \\
&\qquad \text{for } j = 1, 2, \cdots, (2m+1) < +\infty, \ q^{(2)}{}_{jj} = +1 \text{ for } j > 2m+1, \\
(124)\ &\mathrm{III}: \ Q \sim Q^{(3)} = \| q^{(3)}{}_{jk} \|, \ q^{(3)}{}_{jk} = 0 \ j \neq k, \ q^{(3)}{}_{jj} = +1 \\
&\qquad \text{for } j = 1, 2, \cdots, m < +\infty, \ q^{(3)}{}_{jj} = -1 \text{ for } j > m,
\end{aligned}$$

$$\text{IV}: \quad Q \sim Q^{(4)} = \| q^{(4)}{}_{jk} \|, \; q^{(4)}{}_{jk} = 0 \;\; j \neq k, \; q^{(4)}{}_{jj} = q_{j+1\,j+1} = -1$$
$$\text{for } j = 1, 5, 9, \cdots, \; q^{(4)}{}_{jj} = q^{(4)}{}_{j+1\,j+1} = +1 \;\; \text{for } j = 3, 7, 11, \cdots,$$

On pages 607-612, we have shown that simultaneously with the transformation of the matrix B of Theorem I to a given canonical form (87) and hence simultaneously with the reduction of the matrix Q, defined by (123), to a canonical form in (124), the skew-symmetric matrix S of Theorem I is transformed into one of the four following forms:

$$\text{I}: \quad S \sim S^{(1)} = \left\| \begin{array}{c|c} S_{2m} & \\ \hline & \end{array} \right\|, \quad S_{2m} = \bar{S}_{2m} = -S'_{2m},$$

$$\text{II}: \quad S \sim S^{(2)} = \left\| \begin{array}{c|c} S_{2m+1} & \\ \hline & \end{array} \right\|, \quad S_{2m+1} = \bar{S}_{2m+1} = -S'_{2m+1},$$

$$\text{III}: \quad S \sim S^{(3)} = \left\| \begin{array}{c|c} & \\ \hline & S_{\infty-m} \end{array} \right\|, \quad S_{\infty-m} = \bar{S}_{\infty-m} = -S'_{\infty-m},$$

$$(125) \qquad \text{IV}: \quad S \sim S^{(4)} = \left\| \begin{array}{ccccc} S_{11} & Z & S_{13} & Z & \cdots \\ Z & Z & Z & Z & \cdots \\ S_{31} & Z & S_{33} & Z & \cdots \\ Z & Z & Z & Z & \cdots \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \end{array} \right\|,$$

$$S_{pq} = \bar{S}_{pq} = -S'_{qp} = \left\| \begin{array}{cc} \alpha_{pq} & \beta_{pq} \\ \gamma_{pq} & \delta_{pq} \end{array} \right\|, \quad Z = \left\| \begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array} \right\|.$$

[cf. for example for I of (125) formulas (91), (92), and (97)]. Since, according to Theorem IV, we may express any given orthogonal matrix O uniquely in the form (120) it is possible to effect a division of the orthogonal matrices into four classes, namely according as

$$(126) \qquad \text{Class I}: \; O \sim Q^{(1)} e^{-S^{(1)}}, \qquad \text{Class II}: \; O \sim Q^{(2)} e^{-S^{(2)}},$$
$$\text{Class III}: \; O \sim Q^{(3)} e^{-S^{(3)}}, \qquad \text{Class IV}: \; O \sim Q^{(4)} e^{-S^{(4)}},$$

in which the matrices $Q^{(i)}$, $S^{(i)}$ $(i = 1, 2, 3, 4)$ are defined in (124) and (125). As has been mentioned in connection with Theorem I this classification is possible because of the invariance of the spectra and the "covariance" property there pointed out with respect to orthogonal matrix transformations.

We now confine our attention to orthogonal matrices of Classes I and II in (126) and obtain a further simplification of the canonical forms in

(126) for these two classes. We shall understand by a canonical l-rowed skew-symmetric matrix $\boldsymbol{F}_l$ $(l < +\infty)$ the skew-symmetric matrix obtained by "inserting" in succession the real binary skew-symmetric matrices

$$(127)\quad \boldsymbol{S}_{pp} = \left\| \begin{matrix} 0 & s_p \\ -s_p & 0 \end{matrix} \right\|, \; 1 \leqq p \leqq l/2 \;(l \text{ even});\; 1 \leqq p \leqq (l-1)/2 \;(l \text{ odd} \quad -\infty < s_p < +\infty$$

in the diagonal positions of the l-rowed square zero matrix. For even l every row and column of $\boldsymbol{F}_l$ contains one and only one s_l, the same being true for odd l with the exception of the first row and column which contain only zeros. As is well known every finite real skew-symmetric matrix $\boldsymbol{S}$ is orthogonally equivalent to a canonical skew-symmetric matrix so that the points in the spectrum of $\boldsymbol{S}$ are then necessarily

$$(128)\quad \pm is_1, \pm is_2, \cdots, \pm is_{l/2} \quad \text{or} \quad 0, \pm is_1, \pm is_2, \cdots, is_{(l-1)/2},$$

according as l is even or odd respectively. In an analogous manner we shall understand by an infinite canonical skew-symmetric matrix F_∞ the skew-symmetric matrix obtained by inserting infinitely many of the binary skew-symmetric matrices (127) in the diagonal positions of the infinite zero matrix. Those bounded skew-symmetric matrices, possessing no continuous spectrum, are analogous to finite real skew-symmetric matrices in that they are orthogonally equivalent † to an infinite canonical skew-symmetric matrix F_∞. With the use of canonical skew-symmetric matrices, as defined above, the canonical forms for the matrices of Classes I and II as set forth in (126) may be further simplified to

$$(129)\quad \text{Class I}: \mathrm{O} \sim \mathrm{Q}^{(1)}e^{-\mathrm{S}^{(1)}} \sim \mathrm{Q}^{(1)}e^{\mathrm{F}_{2m}}, \quad \text{Class II}: \mathrm{O} \sim \mathrm{Q}^{(2)}e^{-\mathrm{S}^{(2)}} \sim \mathrm{Q}^{(2)}e^{\mathrm{F}_{2m+1}},$$

inasmuch as, as is clear from I and II of (126) that

$$\mathrm{Q}^{(1)}e^{-\mathrm{S}^{(1)}} = \mathrm{Q}^{(1)}e^{-\mathrm{S}_{2m}}, \qquad \mathrm{Q}^{(2)}e^{-\mathrm{S}^{(2)}} = \mathrm{Q}^{(2)}e^{-\mathrm{S}_{2m+1}}.$$

By employing the identity ‡

† See "Hi", p. 163, where the above fact is demonstrated for completely continuous skew-symmetric matrices the proof permitting an immediate generalization to bounded skew-symmetric matrices, not possessing continuous spectra.

‡ The identity (130) may be derived on applying the Cayley-Hamilton theorem to the definition of $\exp \mathrm{M}$ as given on p. 588. On placing

$$\boldsymbol{F}_\theta = \left\| \begin{matrix} 0 & \theta \\ -\theta & 0 \end{matrix} \right\|,$$

the Cayley-Hamilton theorem yields the recursion formulas

$$(130) \qquad \exp \begin{Vmatrix} 0 & \theta \\ -\theta & 0 \end{Vmatrix} = \begin{Vmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{Vmatrix}$$

and if l is odd also

$$\exp \begin{Vmatrix} 0 & 0 & 0 \\ 0 & 0 & \theta \\ 0 & -\theta & 0 \end{Vmatrix} = \begin{Vmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{Vmatrix},$$

in which θ denotes an arbitrary real or complex number, and the finite canonical l-rowed rotation matrices $\boldsymbol{O}_l$ defined by

$$(131) \qquad \boldsymbol{O}_l = e^{\mathrm{F}_l},$$

we are enabled to write (129) in the form

$$(132) \qquad \text{Class I:}\quad \mathrm{O} \sim \mathrm{Q}^{(1)}\boldsymbol{O}_{2m}, \qquad \text{Class II:}\quad \mathrm{O} \sim Q^{(2)}\boldsymbol{O}_{2m+1},$$

and consequently replace the classification of the orthogonal matrices O as given in (126) by

$$(133) \qquad \begin{aligned} \text{Class I:}&\quad \mathrm{O} \sim \mathrm{Q}^{(1)}e^{-\mathrm{S}^{(1)}} \sim \mathrm{Q}^{(1)}\boldsymbol{O}_{2m}, \\ \text{Class II:}&\quad \mathrm{O} \sim \mathrm{Q}^{(2)}e^{-\mathrm{S}^{(2)}} \sim \mathrm{Q}^{(2)}\boldsymbol{O}_{2m+1}, \\ \text{Class III:}&\quad \mathrm{O} \sim \mathrm{Q}^{(3)}e^{-\mathrm{S}^{(3)}}, \\ \text{Class IV:}&\quad \mathrm{O} \sim \mathrm{Q}^{(4)}e^{-\mathrm{S}^{(4)}}. \end{aligned}$$

For finite orthogonal matrices there exist no classes which would correspond to Classes III and IV of (133) whereas Classes I and II correspond to the finite orthogonal matrices of determinants $+1$ and -1 respectively since, denoting by $\det \boldsymbol{M}$ the determinant of an arbitrary finite matrix $\boldsymbol{M}$, we have from † (124)

$$\boldsymbol{F}_\theta^{2j} = (-1)^j\theta^{2j}\boldsymbol{E}, \quad \boldsymbol{F}_\theta^{2j+1} = (-1)^j\theta^{2j}\boldsymbol{F}_\theta; \qquad (j = 0, 1, 2, \cdots),$$

from which the matrix

$$e^{\mathrm{F}\theta} = \sum_{\nu=0}^{+\infty} \frac{\boldsymbol{F}_\theta^{\nu}}{\nu!} = \sum_{\nu=0}^{+\infty} \frac{\boldsymbol{F}_\theta^{2\nu}}{(2\nu)!} + \sum_{\nu=0}^{+\infty} \frac{\boldsymbol{F}_\theta^{2\nu+1}}{(2\nu+1)!},$$

takes the form

$$e^{\mathrm{F}\theta} = \boldsymbol{E}\sum_{\nu=0}^{+\infty} (-1)^\nu \frac{\theta^{2\nu}}{(2\nu)!} + \frac{\boldsymbol{F}_\theta}{\theta}\sum_{\nu=0}^{+\infty} (-1)^\nu \frac{\theta^{2\nu+1}}{(2\nu+1)!}$$

$$= \boldsymbol{E}\cos\theta + \frac{\sin\theta}{\theta}\boldsymbol{F}_\theta = \begin{Vmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{Vmatrix}.$$

Cf. the paper of Galanti referred to on p. 594.

† The first of the equations in (134) follows from the invariance of a determinant of a matrix with respect to matrix transformations, i. e.

$$\det e^{-\mathrm{S}} = \det \boldsymbol{O}e^{-\mathrm{S}}\boldsymbol{O}^{-1} = \det e^{-\boldsymbol{O}\mathrm{S}\boldsymbol{O}^{-1}},$$

and the easily verified fact that the finite canonical rotation matrices defined in (131) above are of determinant $+1$.

$$
\begin{aligned}
&\det e^{-S} = (+1), \qquad S = \bar{S} = -S',\\
(134)\qquad &\det Q^{(1)} = (-1)^{2m} = +1,\\
&\det Q^{(2)} = (-1)^{2m+1} = -1.
\end{aligned}
$$

VI. Rotations and Reflections.

We now return to infinite orthogonal matrices and propose the definition: *We designate an orthogonal matrix* O *as a rotation matrix if and only if there exists a bounded real skew-symmetric matrix* G *(not necessarily unique) for which*

$$(135)\qquad \mathrm{O} = e^{\mathrm{G}},$$

holds. An orthogonal matrix not permitting this representation we designate as a reflection matrix.

In order to justify the introduction of this definition we show † the following lemma pertaining to finite orthogonal matrices:

Lemma III. *Every orthogonal matrix of determinant* $+1$ *(i. e. a finite rotation matrix) permits the representation* (135) *and no orthogonal matrix of determinant* -1 *(i. e. a finite reflection matrix) permits this representation.*

Let $\boldsymbol{O}$ be a given orthogonal matrix of determinant $+1$, i. e. the matrix $\boldsymbol{O}$ is of Class I in (133) so that we have

$$(136)\qquad \boldsymbol{O} \sim \boldsymbol{Q}^{(1)} e^{\mathrm{F}_{2m}},$$

in which the matrices $\boldsymbol{Q}^{(1)}$ and $\boldsymbol{F}_{2m}$ are defined on p. 616 and p. 618, respectively. If we now introduce a canonical skew-symmetric matrix $\boldsymbol{F}^{(1)}{}_{2m}$ defined by placing

† That every orthogonal matrix of determinant $+1$ permits the representation (135) is not, as appears at first sight, a trivial consequence of the normal forms for orthogonal matrices under orthogonal matrix transformations and the above identity (130) is clear from the example

$$
\begin{Vmatrix}
\dfrac{\pi^2-2b^2}{\pi^2} & -\dfrac{2b(\pi^2-b^2)^{1/2}}{\pi^2} & 0\\[2ex]
-\dfrac{2b(\pi^2-b^2)^{1/2}}{\pi^2} & -\dfrac{\pi^2-2b^2}{\pi^2} & 0\\[2ex]
0 & 0 & -1
\end{Vmatrix}
= \exp
\begin{Vmatrix}
0 & 0 & b\\
0 & 0 & (\pi^2-b^2)^{1/2}\\
-b & -(\pi^2-b^2)^{1/2} & 0
\end{Vmatrix}
$$

of an orthogonal matrix of determinant $+1$ which is already in canonical form and whose exponential representation (135) is not at all obvious from the identity (130).

$$s_i = 1 \qquad (i = 1, 2, \cdots),$$

in (128) we are enabled to write the matrix $\boldsymbol{Q}^{(1)}$ of (136), (124) in the exponential form

$$(137) \qquad \boldsymbol{Q}^{(1)} = e^{\pi F^{(1)}_{2m}},$$

inasmuch as, as follows from (130) for $\theta = \pi$, we have

$$\exp \pi \begin{Vmatrix} 0 & 1 \\ -1 & 0 \end{Vmatrix} = \begin{Vmatrix} -1 & 0 \\ 0 & -1 \end{Vmatrix}.$$

Since the matrices $\boldsymbol{F}_{2m}$ and $\boldsymbol{F}^{(1)}{}_{2m}$ are obviously commutable we may employ the distributive property of the exponential representation with respect to commutable matrices and write (136) with the help of (137) in the form

$$\boldsymbol{O} \sim e^{\pi F^{(1)}_{2m}} \cdot e^{F_{2m}} = \exp\,[\pi \boldsymbol{F}^{(1)}{}_{2m} + \boldsymbol{F}_{2m}],$$

and hence, since the property of being a rotation matrix in the sense of our definition is for finite as well as for infinite matrices obviously an orthogonal invariant,† the demonstration of the first part of Lemma III is completed.

The demonstration of the second part of Lemma III, namely that every finite matrix permitting the representation (133) is necessarily of determinant $+1$ is contained in (134) (cf. footnote thereto on page 619).

As an application of the exponential definition of rotation and reflection matrices as given in (135) we show

THEOREM V. *Every orthogonal matrix of Class I is, in the sense of the definition* (135), *a rotation matrix and every orthogonal matrix of Class II is, in the sense of definition* (135), *a reflection matrix.*

The demonstration of the first part of this theorem is seen to be identical, from the formula (136) on, of the first part of Lemma III and consequently is not repeated. As has been mentioned above, the property of an orthogonal matrix to be a rotation or a reflection matrix, in the sense of the definition on p. 620, is an orthogonal invariant. Consequently in order to demonstrate that the matrices of Class II in (133) are all reflection matrices it will be sufficient to demonstrate that the matrices $Q^{(2)}\boldsymbol{O}_{2m+1}$ of II in (133) are all reflection matrices. The second part of Theorem V may not, as appears at

---

† This invariance property is trivial inasmuch as follows from the definition of e^G on p. 588 that if O is an arbitrary orthogonal matrix then $Oe^GO^{-1} = e^{OGO^{-1}}$ in which OGO^{-1} is again a bounded real skew-symmetric matrix.

first sight, be reduced to the fact that no finite orthogonal matrix of determinant -1 permits the representation (135) [cf. (130)], since it would be possible that $Q^{(2)}e^{F_{2m+1}} = Q^{(2)}O_{2m+1} = e^{G}$ in which G is a bounded real skew-symmetric matrix containing infinitely many elements different from zero. In order to give a correct demonstration we first show

LEMMA IV. *The spectrum of an arbitrary real bounded Hermitian matrix* H (*not necessarily reduced*) *is mapped upon the spectrum of the matrix* e^{iH} *by the transformation* $w = e^{iz}$, *i. e. if* θ (θ *real*) *is a point in the spectrum of* H *then* $e^{i\theta}$ *is a point in the spectrum of* e^{iH}.

The proof of this lemma is trivial for reduced Hermitian matrices as it follows directly from the connection between the spectral matrices of H and e^{iH} as given in (59). We shall however require the result of this lemma for an arbitrary Hermitian matrix H and accordingly give a direct proof valid even if the matrix is not reduced. We adopt the notations for an arbitrary Hermitian matrix M

$$(138)\qquad \begin{array}{lll} \boldsymbol{l}\,[M] & \text{greatest lower bound of} \quad \Phi\,[M;x] & \text{for} \quad |x| = 1, \\ \boldsymbol{u}\,[M] & \text{least upper bound of} \quad |\Phi\,[M;x]| & \text{for} \quad |x| = 1. \end{array}$$

We shall require the well known inequalities †

$$(139)\qquad |\Phi\,[MN;x]| \leqq (\Phi\,[MM^*;x]\,\Phi\,[N^*N;x])^{1/2} \quad \text{for} \quad |x| = 1,$$

$$(140)\qquad \boldsymbol{u}\,[M,N] \leqq \boldsymbol{u}\,[M]\,\boldsymbol{u}\,[N],$$

$$(141)\qquad \boldsymbol{l}\,[MM^*] \geqq 0 \quad \boldsymbol{l}\,[M^*M] \geqq 0,$$

valid for any two bounded matrices M and N.

First we derive the relations

$$(142)\qquad \Phi[(\theta E - H)(\theta E - H)^*;x] = \Phi[(\theta E - H)^*(\theta E - H);x] = \Phi[(\theta E - H)^2;x] \geqq 0,$$

$$(143)\qquad \Phi[(e^{i\theta}E - e^{iH})(e^{i\theta}E - e^{iH})^*;x] = \Phi[(e^{i\theta}E - e^{iH})^*(e^{i\theta}E - e^{iH});x] = -2\sum_{\nu=1}^{+\infty}(-1)^{\nu}\frac{\Phi[(\theta E - H)^{2\nu};x]}{(2\nu)!}$$

$$(144)\qquad 0 \leqq \Phi[(\theta E - H)^{2k};x] \leqq (\Phi[(\theta E - H)^2;x]\,\Phi[(\theta E - H)^{4k-2};x])^{1/2}; \quad (k = 1, 2, \cdots),$$

valid for real θ, an arbitrary unit vector x, and an arbitrary bounded Hermitian matrix H. The proof of (142) is, from (141), trivial. For the demonstration of (143) we observe that from

† Cf. for instance "S", p. 137, p. 130, p. 132, respectively.

$$(145)\quad \Phi[(e^{i\theta}E-e^{iH})(e^{i\theta}E-e^{iH})^*;x] \\ =\Phi[(e^{i\theta}E-e^{iH})(e^{-i\theta}E-e^{-iH});x] \\ =\Phi[(e^{i\theta}E-e^{iH})^*(e^{i\theta}E-e^{iH});x],$$

we obtain

$$(146)\quad \Phi[(e^{i\theta}E-e^{iH})(e^{i\theta}E-e^{iH})^*;x]=2-\Phi[(e^{i(\theta E-H)}+e^{-i(\theta E-H)});x],$$

from which (143) obviously follows. The inequalities (144) are derived from that of (139) by placing therein $M=(\theta E-H)$, $N=(\theta E-H)^{2k-1}$ and $k=2,3\cdots$ in succession and at the same time employing (142) and

$$0\leqq\Phi[(\theta E-H)^2(\theta E-H)^{2*};x]=\Phi[(\theta E-H)^4;x].$$

If H is a diagonal matrix the lemma is trivial and we exclude this case in what follows. Let θ be a point in the spectrum of H, i. e. assume [cf. (581)]

$$(147)\quad \boldsymbol{l}[(\theta E-H)(\theta E-H)^*] \\ =\boldsymbol{l}[(\theta E-H)^*(\theta E-H)]=\boldsymbol{l}[(\theta E-H)^2]=0,$$

so that there exists a sequence of unit vectors $\{x_n\}$ for which

$$(148)\quad \lim_{n\to+\infty}\Phi[(\theta E-H)^2;x_n]=0.$$

From (142), (143), and (144) we have

$$(149)\quad 0\leqq\Phi[(e^{i\theta}E-e^{iH})(e^{i\theta}E-e^{iH})^*;x]\leqq 2\sum_{\nu=1}^{+\infty}\frac{\Phi[(\theta E-H)^{2\nu},x]}{(2\nu)!} \\ \leqq 2(\Phi[(\theta E-H)^2;x])^{1/2}\sum_{\nu=1}^{+\infty}\frac{(\Phi[(\theta E-H)^{4\nu-2};x])^{1/2}}{(2\nu)!}.$$

Since $\theta E-H\neq\|0\|$ is bounded we have

$$(150)\quad 0<\boldsymbol{u}[\theta E-H]=\boldsymbol{u},$$

where we write, for the sake of brevity, $\boldsymbol{u}[\theta E-H]=\boldsymbol{u}$ and hence have from (138) and (140)

$$(\Phi[(\theta E-H)^{4\nu-2};x])^{1/2}\leqq(\boldsymbol{u}[(\theta E-H)^{4\nu-2}])^{1/2}\leqq\boldsymbol{u}^{2\nu-1};\quad(\nu=1,2,\cdots),$$

so that the inequality (149) may be replaced by

$$(151)\quad 0\leqq\Phi[(e^{i\theta}E-e^{iH})(e^{i\theta}E-e^{iH})^*;x] \\ \leqq\frac{2}{\boldsymbol{u}}(\Phi[(\theta E-H)^2;x])^{1/2}\sum_{\nu=1}^{+\infty}\frac{\boldsymbol{u}^{2\nu}}{(2\nu)!}<\frac{2e^{\boldsymbol{u}}}{\boldsymbol{u}}(\Phi[(\theta E-H)^2;x])^{1/2}.$$

If one now applies the inequality (151) to the sequence of vectors $\{x_n\}$ for which (148) holds there follows from (151) the fulfillment of (3) for

$$\lambda = e^{i\theta}, \qquad M = e^{iH},$$

i. e. $e^{i\theta}$ is a point in the spectrum of e^{iH}.

If S is a real bounded skew-symmetric matrix the matrix iS is a bounded Hermitian matrix so that we have the

COROLLARY. *The spectrum of an arbitrary real bounded skew-symmetric matrix* S *is mapped upon the spectrum of the matrix* e^S *by the transformation* $w = e^z$, *i. e. if* $i\theta$ *is a point in the spectrum of* S *then* $e^{i\theta}$ *is a point in the spectrum of* e^S.

We now demonstrate the second part of Theorem V and first calculate the matrix $Q^{(2)}\boldsymbol{O}_{2m+1}$ of (133) obtaining from (124), (132), (127) and (131)

(152) $Q^{(2)}\boldsymbol{O}_{2m+1}$

$$= \left\| \begin{array}{cccccccccccc} -1 \\ & -\cos s_1 & -\sin s_1 \\ & \sin s_1 & -\cos s_1 \\ & & & -\cos s_2 & -\sin s_2 \\ & & & \sin s_2 & -\cos s_2 \\ & & & & & \ddots \\ & & & & & & -\cos s_m & -\sin s_m \\ & & & & & & \sin s_m & -\cos s_m \\ & & & & & & & & 1 \\ & & & & & & & & & 1 \\ & & & & & & & & & & \ddots \end{array} \right\|,$$

from which it is clear that the matrix (152) possesses only a point spectrum namely the points

$$(153) \qquad -1,\ e^{\pm s_1},\ e^{\pm is_2}, \cdots, e^{\pm is_m},\ 1,\ 1, \cdots$$

so that -1 necessarily occurs an odd number of times in the spectrum of (152). If now the matrix (152) were to permit the exponential representation e^G of (135) the matrix G therein could, from the above corollary, possess no continuous spectra and consequently (cf. p. 618) is orthogonally equivalent to an infinite canonical skew-symmetric matrix F_∞. Since the spectrum of any matrix is invariant under matrix transformation it is clear that either the points

$$(154) \qquad (2k+1)\pi i, \qquad -(2k+1)\pi i; \qquad (k = 0, 1, \cdots),$$

belong simultaneously to the spectrum of G or neither belongs to the spectrum

of G. Therefore if -1 occurs in the spectrum of e^G it follows from (154) and the above corollary that it must occur an even or an infinite number of times which is a contradiction to (153). q. e. d.

APPENDIX.

The Problem of moments for a finite interval and the Theory of Bounded Hermitian Matrices.

Several presentations have been given for the spectral representation of a bounded Hermitian matrix by means of Stieltjes integrals as introduced by Hilbert. The various modes of investigation require in common with the method of Hilbert, as far as the author knows from the literature,† the use of the characteristic constants of the truncated matrices of the given bounded Hermitian matrix. This procedure has certain methodical inconveniences. First, the spectra of the truncated matrices of a bounded Hermitian matrix do not necessarily determine the spectrum of the infinite matrix itself. An example of such a bounded Hermitian matrix has been pointed out by Toeplitz ‡ in which he considers the bounded Hermitian matrix constructed from the infinite zero matrix by inserting the real binary symmetric matrix

$$\begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix}$$

in the diagonal positions of the infinite zero matrix, i. e. the matrix of the Hermitian form

$$x_1\bar{x}_2 + x_2\bar{x}_1 + x_3\bar{x}_4 + x_4\bar{x}_3 + \cdots,$$

and which contains only the points -1 and $+1$ in its spectrum whereas not only -1 and $+1$ but also 0 belong to the spectrum of infinitely many truncated matrices. Second, the method of the truncated matrices is restricted to the treatment of Hermitian matrices alone. For instance the truncated matrices of a unitary matrix are not necessarily also unitary matrices although it is possible to obtain by a direct treatment of the trigonometrical momentum problem belonging to the infinite matrix itself a spectral

† I. See "Hi", pp. 131-137; II. Hellinger Toeplitz, "Integralgleichungen und Gleichungen mit unendlichvielen Unbekannten," *Encyklopädie der Mathematische Wissenschaften* 2III2, pp. 1575-1584. III. F. Riesz, "Über quadratische Formen von unendlichvielen Veränderlichen," *Göttingen Nachrichten* (1910), pp. 190-195. Cf. also by the same author "R", pp. 128-139. We refer to the article of Hellinger Toeplitz as "HT" and this paper of Riesz as "U".

‡ See "T", p. 107.

representation which is entirely analogous to that discovered by Hilbert † for bounded Hermitian matrices. It is shown in the present note that such a direct treatment is also possible for bounded Hermitian matrices and yields a very simple demonstration of the fundamental theorem of Hilbert. It is necessary to start, not with the problem of trigonometrical moments but with the ordinary problem of moments for a finite interval as developed by Hausdorff.‡ The method possesses, in contrast to the methods of F. Riesz, not only the advantage of avoiding the use § of the truncated matrices at all but also enables us to start directly with a monotone Stieltjes kernel (Belegungsfunktion) whereas F. Riesz deduces ¶ the result from his general theorem ‖ on the solvability of a system of Stieltjes integral equations with a function of bounded variation and then has to prove that the solution is actually monotone. It is of course a matter of fact that the demonstration of the theorem of Hausdorff is based on the theorems of Helly applied by Carleman in connection with the method of the truncated matrices as developed by Hilbert. In addition to the theorem of Hausdorff we need the fact that every bounded positive definite matrix possesses a bounded "square root". If one is already in possession of the spectral representation of a bounded Hermitian matrix (the deduction of which is the purpose of this note) the uniqueness,†† as well as the existence of a bounded positive definite "square root" of a bounded positive definite matrix, is assured. We need, however, only the existence of a "square root" of a bounded positive definite matrix and this is available, according to a remark of eHllinger-Toeplitz,‡‡ without the use of a spectral representation.

For the sake of completeness we first discuss in some detail the remark of Hellinger-Toeplitz concerning the construction of a bounded positive definite square root of a bounded positive definite matrix. The basis of this remark is the ordinary Taylor development

† See "Hi", p. 138.

‡ See "Ha".

§ See "R", pp. 129-139, where Riesz gives a direct proof employing an approximation by matrix polynomials which however also depends upon the use of the truncated matrices.

¶ See "U".

‖ F. Riesz, "Sur les opérations fonctionelles linéaires," *Comptes Rendus*, 29 novembre (1909), and "Sur certains systèmes d'equations fonctionelles et l'approximation des fonctions continues," *Comptes Rendus*, 14 mars (1910).

†† See "W", p. 267. For finite matrices a direct demonstration of this existence and uniqueness property is given by L. Autonne in "Sur l'Hermitien," *Rendiconti del Circolo Matematico di Palermo*, Vol. 16 (1902), pp. 120-121.

‡‡ See "HT", p. 1567, footnote (522 a).

(155) $$f(z) = (1-z)^{1/2} = \sum_{\nu=0}^{+\infty} a_\nu z^\nu, \qquad |z| < 1,$$

in which.

(156) $$a_0 a_0 = 1, \quad a_0 a_1 + a_1 a_0 = -1, \quad \sum_{\mu=0}^{n} a_\mu a_{n-\mu} = 0; \qquad (n = 2, 3, \cdots).$$

We denote by

(157) $$\Psi[\mathrm{A}; x, y] = \sum_{p=1}^{+\infty} \sum_{q=1}^{+\infty} a_{pq} x_p y_q,$$

the bilinear form of an arbitrary bounded matrix A and by

(158) $$\Phi[\mathrm{A}; x] = \Psi[\mathrm{A}; x, \bar{x}] = \sum_{p=1}^{+\infty} \sum_{q=1}^{+\infty} a_{pq} x_p \bar{x}_q,$$

the compound form (Kopplungsform) of A. The upper limit of $|\Psi[\mathrm{A}; x, y]|$ for those vectors x and y for which

(159) $$|x| = (\sum_{\nu=1}^{+\infty} |x_\nu|^2)^{1/2} = 1, \qquad |y| = (\sum_{\nu=1}^{+\infty} |y_\nu|^2)^{1/2} = 1,$$

will be denoted by $\boldsymbol{R}(\mathrm{A})$, the upper and lower limits of $|\Phi[\mathrm{A}; x]|$ for those vectors x of (159) being denoted by $\boldsymbol{M}(\mathrm{A})$ and $\boldsymbol{m}(\mathrm{A})$ respectively. It is known that if A_1 and A_2 are two arbitrary bounded matrices then according to the inequality of Schwarz †

(160) $$\boldsymbol{R}(\mathrm{A}_1\mathrm{A}_2) \leqq \boldsymbol{R}(\mathrm{A}_1)\boldsymbol{R}(\mathrm{A}_2).$$

Furthermore it is clear that the series

(161) $$\sum_{\nu=0}^{+\infty} \mathrm{A}_\nu$$

converges to a bounded matrix A if

(162) $$\sum_{\nu=0}^{+\infty} \boldsymbol{R}(\mathrm{A}_\nu) < +\infty.$$

For an arbitrary bounded Hermitian matrix H it is known that the compound form $\Phi[\mathrm{H}; x]$ takes only real values and that

(163) $$\boldsymbol{R}(\mathrm{H}) = \boldsymbol{M}(\mathrm{H}).$$

A bounded Hermitian matrix for which the upper and lower limits of its compound form on the complex Hilbert sphere $|x| = 1$ of (159) are con-

† Cf. for instance "S", p. 130.

tained in the open interval $(0,1)$ will be denoted by K. We accordingly have for every Hermitian matrix K

$$(164)\qquad 0<\boldsymbol{m}(\mathrm{K})\leqq \boldsymbol{M}(\mathrm{K})<1,$$

in which the equality sign holds for and only for those matrices $\gamma\mathrm{E}$ in which E is the unit matrix and $0<\gamma<1$. The matrix $\mathrm{E}-\mathrm{K}$ is likewise a bounded Hermitian matrix for which we have from (164)

$$(165)\qquad 0<\boldsymbol{m}(\mathrm{E}-\mathrm{K})\leqq \boldsymbol{M}(\mathrm{E}-\mathrm{K})<1.$$

The matrix

$$(166)\qquad \mathrm{A}=\sum_{\nu=0}^{+\infty} a_\nu K^\nu,$$

in which the a_ν are the coefficients in the Taylor development (155) and consequently fulfill the bilinear relations (156), is a bounded matrix since, from (164) and the definition of the a_ν in (155), (156) we have from (160) and (163)

$$\sum_{\nu=0}^{+\infty}\boldsymbol{R}(a_\nu K^\nu)=\sum_{\nu=0}^{+\infty}|\,a_\nu\,|\,\boldsymbol{R}(K^\nu)\leqq\sum_{\nu=0}^{+\infty}|\,a_\nu\,|\,\boldsymbol{R}(K)^\nu=\sum_{\nu=0}^{+\infty}|\,a_\nu\,|\,\boldsymbol{M}(K)^\nu<+\infty,$$

and therefore

$$\boldsymbol{R}(\mathrm{A})\leqq\sum_{\nu=0}^{+\infty}\boldsymbol{R}(a_\nu\mathrm{K}^\nu)<+\infty.$$

One then verifies easily from (166) and the definition of the a_ν in (155), (156) that

$$(167)\qquad \mathrm{A}^2=\mathrm{E}-\mathrm{K},$$

and, in an analogous manner from (165), that the matrix

$$(168)\qquad \mathrm{B}=\sum_{\nu=0}^{+\infty} a_\nu(\mathrm{E}-\mathrm{K})^\nu,$$

is bounded and fulfills the matrix equation

$$(169)\qquad \mathrm{B}^2=\mathrm{K}.$$

The matrices A and B defined in (166) and (168) may then be said to be "square roots" of the bounded positive definite matrices $\mathrm{E}-\mathrm{K}$ and K respectively. The matrices A and B in (166) and (168) are in addition positive definite but inasmuch as we shall not need this fact in the following we give no demonstration. The matrices A and B are obviously Hermitian and commutable, i. e.

$$(170)\qquad \mathrm{AB}=\mathrm{BA}=\mathrm{A}^*\mathrm{B}^*=\mathrm{B}^*\mathrm{A}^*\quad\text{where}\quad \mathrm{A}^*=\bar{\mathrm{A}}',\ \mathrm{B}^*=\bar{\mathrm{B}}'.$$

We now employ the results of the previous paragraph in the demonstration † of the

LEMMA. *If* K *is any bounded Hermitian matrix whose compound form* $\Phi[K;x]$ *is bounded by the inequalities* (164), *i. e.*

$$(171) \qquad 0<\theta_1<\Phi[K;x]<\theta_2<1 \quad \text{for} \quad |x|=1,$$

then the matrix polynomials

$$(172) \qquad \pi_{n,m}(K) \quad \text{where} \quad \pi_{n,m}(x)=x^m(1-x)^{n-m}; \\ (m=0,1,\cdots,n);\ (n=0,1,\cdots),$$

are not-negative definite matrices, i. e.

$$(173) \qquad 0\leqq\Phi[\pi_{m,n}(K);x] \quad \text{for} \quad |x|=1; \\ (m=0,1,\cdots,n);\ (n=0,1,\cdots).$$

In order to demonstrate (173) we employ the "square root" representation (167), (169) for the matrices $E-K$ and K respectively in the matrix polynomials (172) to obtain with the aid of the multiplication theorems (Faltungsätze) of Hilbert and the commutability property (170)

$$\pi_{m,n}(K)=(B^2)^m(A^2)^{n-m}=(B^mA^{n-m})(B^mA^{n-m})=(B^mA^{n-m})(B^mA^{n-m})^*; \\ (m=0,1,\cdots,n;\ n=0,1,\cdots),$$

from which (173) follows at once inasmuch as the Hermitian matrix $(B^mA^{n-m})(B^mA^{n-m})^*$ is known to be non-negative definite.

We now consider the problem of moments ‡ for a finite interval, viz.

$$(174) \qquad c_l=\int_0^1 \mu^l d\rho(\mu); \qquad (l=0,1,\cdots),$$

† The demonstration of this lemma may also be seen as a consequence of a theorem given by F. Riesz in "R", p. 129, which states that for a Hermitian matrix H whose compound form satisfies the inequalities

$$m\leqq\Phi[H;x]\leqq M \quad \text{for} \quad |x|=1,$$

the compound form of the matrix polynomial

$$f(H) \quad \text{where} \quad f(x)=\sum_{\nu=0}^{n} c_\nu x^\nu,$$

constructed from the matrix H, satisfies the inequalities

$$\lambda\leqq\Phi[f(H);x]\leqq\Lambda \quad \text{for} \quad |x|=1,$$

where

$$\lambda\leqq f(\mu)\leqq\Lambda \quad \text{for} \quad m\leqq\mu\leqq M.$$

In the demonstration for this theorem, given by Riesz, use is made, however, of the truncated matrices of the matrix $f(H)$ whereas in the direct demonstration of the lemma given above, we avoid this procedure.

‡ See "Ha", pp. 222-232.

in which the c_l are preassigned real numbers and there is desired a real monotone non-decreasing function $\rho(\mu)$ of the real argument μ which satisfies the system (174) of Stieltjes integral equations. According to Hausdorff ‡ this problem possesses a solution then and only then if the sequence $\{c_l\}$ is completely monotone, i. e. if and only if the following inequalities hold:

$$(175)\qquad \Delta^{n-m}c_m \geqq 0; \qquad (m = 0, 1, \cdots, n); \; (n = 0, 1, \cdots),$$

where

$$\Delta c_m = c_m - c_{m+1}, \qquad \Delta^j = \Delta^{j-1}\Delta; \qquad (j = 2, 3, \cdots),$$

so that

$$\Delta^{n-m}c_m = c_m - \binom{n-m}{1}c_{m+1} + \cdots + (-1)^{n-m}c_n;$$
$$(m = 0, 1, \cdots, n); \; (n = 0, 1, \cdots).$$

As is well known one may then insure the uniqueness of the solution $\rho(\mu)$ of (174) at all points of its interval of definition, the closed interval $[0, 1]$, by prescribing, for instance, the normalizing conditions

$$(176)\qquad 0 = \rho(0) \quad \text{and} \quad \rho(\mu - 0) = \rho(\mu), \qquad 0 < \mu < 1,$$

for the solution $\rho(\mu)$ of (174), obtaining as a consequence of the assumption $\rho(0) = 0$ from (174) for $l = 0$

$$(177)\qquad \rho(1) - \rho(0) = \rho(1) = c_0.$$

For our purposes we consider the problem of moments arising from (174) if one places

$$(178)\qquad c_l = \Phi[\mathrm{K}^l; x], \quad |x| = 1, \quad \mathrm{K}^0 = \mathrm{E}; \qquad (l = 0, 1, \cdots),$$

in which x is an arbitrary but fixed unit vector and K a Hermitian matrix whose compound form $\Phi[\mathrm{K}; x]$ is bounded by the inequalities (164), i. e. (171). We seek a monotone non-decreasing function $\rho(\mu; x)$ of the real argument μ and the components $x_1, x_2, \cdots$ of the above given unit vector x which satisfies the system (174), (178) of Stieltjes integral equations, i. e. is a solution of the problem of moments

$$(179)\qquad \Phi[\mathrm{K}^l; x] = \int_0^1 \mu^l \, d\rho(\mu; x), \quad |x| = 1; \qquad (l = 0, 1, \cdots).$$

That the moment problem (179) possesses, for any given unit vector x, such

‡ See "Ha", p. 226. Cf. also I. J. Schoenberg, "On Finite and Infinite Completely Monotonic Sequences," *Bulletin of the American Mathematical Society*, Vol. 38, No. 2 (1932), pp. 72-76.

a monotone solution follows from the above lemma inasmuch as the Hausdorff conditions (175) are, for the moment problem (179)

$$\Delta^{n-m}\Phi[K^m;x] \geqq 0, \quad \text{where} \quad \Delta^{n-m}\Phi[K^m;x]$$
$$\equiv \Phi[(K^m - \tbinom{n-m}{1}K^{m+1} + \cdots + (-1)^{n-m}K^n);x] = \Phi[K^m(E-K)^{n-m};x],$$

and which are, from (172) and (173), obviously fulfilled. In place of the normalizing conditions (176) prescribed for the solution $\rho(\mu)$ of (174) we now prescribe

(180) $\quad 0 \equiv \rho(0;x)$ and $\rho(\mu - 0;x) \equiv \rho(\mu;x)$ for $0 < \mu < 1$; $|x| = 1$,

for the solution $\rho(\mu;x)$ of (179) obtaining in place of (177)

(181) $$\rho(1;x) - \rho(0;x) \equiv \rho(1;x) \equiv 1 \quad \text{for} \quad |x| = 1,$$

the solution $\rho(\mu;x)$ of (178) being thereby uniquely determined, for any given unit vector x, at all points of the closed interval $[0,1]$.

We now define the function $\rho(\mu;x)$ also for $-\infty < \mu < 0$ and $1 < \mu < +\infty$ taking it identically equal to 0 and $+1$ in the two ranges respectively. With this understanding we write in place of (178)

(182) $$\Phi[K^l;x] = \int_{-\infty}^{+\infty} \mu^l d\rho(\mu;x); \qquad (l = 0, 1, \cdots).$$

Let H be an arbitrary bounded Hermitian matrix. Since H is bounded there exist two real numbers α $(\alpha > 0)$ and β so that the matrix

(183) $$K = \alpha H + \beta E$$

fulfills the condition (164). In denoting by $\rho(\mu;x)$ the solution of the moment problem (179), (180), (181) we introduce a function $\sigma(\mu;x)$ by means of the definition

(184) $$\sigma(\mu;x) \equiv \rho(\alpha\mu + \beta;x), \quad \alpha > 0, \quad -\infty < \mu < +\infty.$$

From (182), (183) and (184) we then have

$$\Phi[H^l;x] = \int_{-\infty}^{+\infty} \mu^l d\sigma(\mu;x),$$

i. e. the spectral representation of Hilbert.

A NOTE ON THE CONVERGENCE OF THE SUCCESSIVE APPROXIMATIONS TO THE SOLUTION OF AN ORDINARY DIFFERENTIAL EQUATION.

By E. K. HAVILAND.

If $f(x, y)$ is continuous in x and y together in $|x - x_0| \leqq a$, $|y - y_0| \leqq b$, where we may assume without restriction that $x_0 = y_0 = 0$, Peano has proved * the existence of at least one solution of the differential equation

$$dy/dx = f(x, y)$$

satisfying the initial condition $y(0) = 0$.

Although there will in general be more than one solution under these conditions, O. Perron,† by a slight modification of the proof of M. Nagumo,‡ has shown that a sufficient condition for the uniqueness of the solution is

$$(1) \qquad |f(x, y') - f(x, y'')| \leqq (|k|/x) |y' - y''|$$

where

$$(2) \qquad |k| \leqq 1,$$

which is seen to be a generalization of the Lipschitz condition. Furthermore, Perron has demonstrated † that no greater value of $|k|$ is sufficient.

Correspondingly, the continuity condition assumed in Peano's proof is not in general sufficient to insure the convergence of the successive approximations to the solution given by

$$y_n(x) = \int_0^x f(t, y_{n-1}(t))\,dt, \qquad y_0(x) \equiv 0$$

M. Müller being the first to give § an example where the successive approximations diverge although $f(x, y)$ is continuous.

---

* See, for example, E. Kamke, *Differentialgleichungen* (Leipzig, 1930), pp. 59-66.

† O. Perron, "Eine hinreichende Bedingung für die Unität der Lösung von Differentialgleichungen erster Ordnung," *Mathematische Zeitschrift*, Vol. 28 (1928), pp. 216-219.

‡ M. Nagumo, "Eine hinreichende Bedingung für die Unität der Lösung von Differentialgleichungen erster Ordnung," *Japanese Journal of Mathematics*, Vol. 3 (1926), pp. 107-112.

§ M. Müller, "Über das Fundamentaltheorem in der Theorie der gewöhnlichen Differentialgleichungen," *Mathematische Zeitschrift*, Vol. 26 (1927), pp. 619-645, especially p. 629.

A Rosenblatt has shown,* however, that (1) is a sufficient condition for the convergence of the successive approximations if condition (2) is replaced by $|k| < 1$. It is accordingly of interest to consider the possibility of the occurrence of larger values of $|k|$ in a sufficient condition for the convergence of the successive approximations. In this note, it is proposed to show by a refinement of Müller's method that condition (1) is not sufficient for convergence when $|k| > 1$. As similar methods have failed to establish either the sufficiency or the non-sufficiency of (1) if $|k| = 1$, this boundary case remains unsolved.

The non-sufficiency of (1) for the convergence of the successive approximations when $|k| > 1$ is established by the following example:

Let $dy/dx = f(x, y)$, where

$$f(x,y) = \begin{cases} 0 & x = 0, & -\infty < y < \infty \\ x^\epsilon & 0 < x \leqq a, & -\infty < y < 0 \\ x^\epsilon - (1+\epsilon)y/x, & 0 < x \leqq a, & 0 \leqq y \leqq |x|^{1+\epsilon} \\ -\epsilon x^\epsilon & 0 < x \leqq a, & |x|^{1+\epsilon} < y < \infty \end{cases}$$

and let $x_0 = y_0 = 0$.

Then

$$y_1(x) = \int_0^x f(t, y_0)\,dt = \int_0^x f(t, 0)\,dt = \int_0^x t^\epsilon dt = x^{1+\epsilon}/(1+\epsilon)$$

$$y_2(x) = \int_0^x f(t, y_1(t))\,dt = \int_0^x f(t, t^{1+\epsilon}/(1+\epsilon))\,dt = 0$$

and so in general

$$\left.\begin{aligned} y_{2k-1}(x) &= x^{1+\epsilon}/(1+\epsilon) \\ y_{2k}(x) &= 0 \end{aligned}\right\} \qquad (k = 1, 2, 3, \cdots).$$

Thus the successive approximations diverge. Furthermore, the limits of the convergent sequences $\{y_{2k-1}(x)\}$, $\{y_{2k}(x)\}$ do not satisfy the differential equation.

Whether or not (1) is a sufficient condition for the convergence of the successive approximations when $|k| = 1$ is a question we are not at present able to answer. In the example $dy/dx = f(x, y)$, where

$$f(x,y) = \begin{cases} 0, & x = 0, & -\infty < y < \infty \\ \tfrac{1}{2}x^\epsilon, & 0 < x \leqq a, & -\infty < y < 0 \\ \tfrac{1}{2}x^\epsilon - y/x, & 0 < x \leqq a, & 0 \leqq y \leqq |x|^{1+\epsilon} \\ -\tfrac{1}{2}x^\epsilon, & 0 < x \leqq a, & |x|^{1+\epsilon} < y < \infty \end{cases}$$

and $x_0 = y_0 = 0$,

---

* A. Rosenblatt, "Über die Existenz von Integralen gewöhnlicher Differentialgleichungen," *Arkiv för Matematik, Astronomi och Fysik*, Vol. 5, No. 2 (1909), pp. 1-4.

the successive approximations may be shown to converge to the solution

$$y = \tfrac{1}{2}\,(1/(2+\epsilon))x^{1+\epsilon}$$

however small the value of ϵ. In the limiting case, where $\epsilon = 0$, the successive approximations do, indeed, diverge, but the corresponding function $f(x,y)$, although bounded, is not continuous at $(0,0)$.

Again, the functions

$$f(x,y) = \begin{cases} 0, & x = 0, \quad -\infty < y < \infty \\ \log(1 \pm x), & 0 < x \leqq a, \quad -\infty < y < 0 \\ \log(1 \pm x) - y/x, & 0 < x \leqq a, \quad 0 \leqq y \leqq |x| \log(1 \pm |x|) \\ 0, & 0 < x \leqq a, \quad |x| \log(1 \pm |x|) < y < \infty \end{cases}$$

lead to converging approximations.

In consequence, there does not appear to be any possibility of setting up a divergent sequence of successive approximations in the case $|k| = 1$ by the method of Müller. It may, of course, be that any search for a function $f(x,y)$ for which the successive approximations diverge under the condition (1) when $|k| = 1$ is futile, for this condition may be sufficient for convergence, though such has not so far been proved to be the case to the best of our knowledge.

Examples such as $dy/dx = y^{1/3}$ show that the convergence of the successive approximations is not sufficient to insure uniqueness of the solution. If an example could be given where the successive approximations diverged when $f(x,y)$ satisfied condition (1) with $|k| = 1$, it would be shown by virtue of the Nagumo-Perron theorem that conversely uniqueness was not a sufficient condition for the convergence of the successive approximations, a question at present undetermined.

THE JOHNS HOPKINS UNIVERSITY.

A THREE-DIMENSIONAL TREATMENT OF GROUPS OF LINEAR TRANSFORMATIONS.

By Deborah May Hickey.

Introduction. In his treatment of the theory of linear transformations Professor L. R. Ford * has introduced the concept of the isometric circle. Used in the theory of discontinuous groups it leads easily to the construction of a fundamental region. The resulting simplification in the treatment of the linear transformation suggested the use of the same notion in connection with space transformations.

The kind of space transformations to be studied is defined as follows. Let Σ_0 be the unit sphere with center at the origin; and let $\sigma_1, \sigma_2, \cdots, \sigma_{2p}$ be an even number of spheres orthogonal to Σ_0. Let inversions be made successively in these spheres. Any point P is carried into a point P_1 by the inversion in σ_1; P_1 into P_2 by the inversion in σ_2; and so on, finally P_{2p-1} into P' by the inversion in σ_{2p}. To each point P there corresponds one and only one point P'. Two such sequences of inversions will be considered to define the same transformation if they result in the same P' for each point P of space.

The effect of the preceding transformation on the surface of Σ_0 is to carry it into itself in a one-to-one and directly conformal manner. Conversely, it will appear that each such transformation of the surface of Σ_0 can be achieved by a unique space transformation of the type defined.

By a stereographic projection of Σ_0 on the complex plane there is set up a one-to-one correspondence between the set of space transformations and all linear transformations.

A study is made of groups of these space transformations. It is found that all discontinuous groups of linear transformations correspond to *properly* discontinuous groups in space. The introduction of the isometric sphere leads to the construction of a fundamental region in a very simple manner. The faces, edges, and vertices of this region have striking properties.

The continuous and directly conformal transformation of the surface of a sphere into itself. If the surface of a sphere be mapped in a one-to-one and continuous manner on itself, there results, in general, a magnification of

* *Automorphic Functions*, McGraw-Hill Book Co. (1929). H. Poincaré was the first to introduce a space transformation corresponding to the linear transformation, *Acta Mathematica*, Vol. 3 (1883-1884), pp. 49-92.

some parts and a diminution of other parts. If the mapping is conformal, we have the following

THEOREM 1. *Let the surface of a sphere be transformed in a one-to-one and directly conformal manner on itself. Then, provided the transformation is not a rotation of the sphere about an axis, the locus of points in the neighborhood of which lengths and areas are unaltered is a small circle.*

By a suitable choice of coördinates we can make Σ_0 the sphere in question. We designate the transformation by S. Let Σ_0 be projected stereographically on a z-plane through its center.* Then, as is well known, the plane undergoes a linear transformation when the sphere is transformed by S. Denote by ds the length of an element of arc on the surface of Σ_0 and by ds' the length of the transformed element. We have

$$ds^2 = d\xi^2 + d\eta^2 + d\zeta^2,$$

where $P(\xi, \eta, \zeta)$ is the initial point of the element.

The point P on Σ_0 and its stereographic projection $z = x + iy$ in the plane are connected by the relations

$$\xi = \frac{2x}{z\bar{z}+1}, \quad \eta = \frac{2y}{z\bar{z}+1}, \quad \zeta = \frac{z\bar{z}-1}{z\bar{z}+1},$$

where $\bar{z}$ is the conjugate imaginary of z; whence

$$ds^2 = \frac{8(dx^2+dy^2)}{(z\bar{z}+1)^2}.$$

Similarly

$$ds'^2 = \frac{8(dx'^2+dy'^2)}{(z'\bar{z}'+1)^2},$$

where $z' = x' + iy'$ is the transform of z by a linear transformation

$$z' = \frac{az+b}{cz+d}, \quad \text{with} \quad ad - bc = 1.$$

Then since $dx'^2 + dy'^2 = dz'd\bar{z}'$ and $dz'/dz = 1/(cz+d)^2$, the condition that $ds' = ds$ is that

$$(z'\bar{z}'+1)^2(cz+d)^2(\bar{c}\bar{z}+\bar{d})^2 = (z\bar{z}+1)^2,$$

which reduces to

$$(1) \qquad (1 - a\bar{a} - c\bar{c})z\bar{z} - (a\bar{b} + c\bar{d})z - (\bar{a}b + \bar{c}d)\bar{z} + (1 - b\bar{b} - d\bar{d}) = 0.$$

* The stereographic projection of the sphere is made by projecting from the point $(0, 0, 1)$ upon the z-plane. The equator of Σ_0 in the z-plane remains fixed.

If $1 - a\bar{a} - c\bar{c} \neq 0$, equation (1) represents a circle C with center and radius respectively

$$p = \frac{\bar{a}b + \bar{c}d}{1 - a\bar{a} - c\bar{c}}, \quad r = (a\bar{a} + b\bar{b} + c\bar{c} + d\bar{d} - 2)^{1/2}/(1 - a\bar{a} - c\bar{c}).$$

The expression under the radical sign is non-negative. Moreover, except for a rotation of Σ_0 into itself, it is positive. In fact,

$$a\bar{a} + b\bar{b} + c\bar{c} + d\bar{d} - 2 = (a - \bar{d})(\bar{a} - d) + (b + \bar{c})(\bar{b} + c) \geqq 0,$$

the equality sign holding if and only if $a = \bar{d}$, $b = -\bar{c}$. For these relations (1) is identically satisfied. They are the conditions for a rotation of Σ_0 about an axis.

If $1 - a\bar{a} - c\bar{c} = 0$ and $1 - b\bar{b} - d\bar{d} \neq 0$, equation (1) reduces to the equation of a straight line, not passing through the origin.

The projection of C back on Σ_0 is a circle I_S. If, in particular, $1 - a\bar{a} - c\bar{c} = 0$ and $1 - b\bar{b} - d\bar{d} \neq 0$, the circle passes through the north pole. Under the transformation S infinitesimal lengths in the neighborhood of I_S are unaltered; hence infinitesimal areas on the surface are unaltered also.

We call I_S the *isometric circle* of the transformation S.

The circle I_S is the complete locus of points in the neighborhood of which lengths on Σ_0 are unaltered by S. For, any point P on Σ_0 not on I_S has as a projection in the plane a point $Q(z_0)$ which is not on C. For z_0 inside C, the left member of (1) is negative, and $ds'^2 > ds^2$; for z_0 outside C, $ds'^2 < ds^2$.

The isometric circle I_S cannot be a great circle of Σ_0. In fact, the distance Δ of the plane of I_S,

$$\begin{aligned}\xi(a\bar{b} + c\bar{d} + \bar{a}b + \bar{c}d) + i\eta(a\bar{b} + c\bar{d} - \bar{a}b - \bar{c}d) \\ + \zeta(a\bar{a} - b\bar{b} + c\bar{c} - d\bar{d}) + (a\bar{a} + b\bar{b} + c\bar{c} + d\bar{d} - 2) = 0,\end{aligned}$$

from the center of Σ_0 is given by

$$\Delta = [(a\bar{a} + b\bar{b} + c\bar{c} + d\bar{d} - 2)/(a\bar{a} + b\bar{b} + c\bar{c} + d\bar{d} + 2)]^{1/2},$$

which is essentially positive, since the numerator vanishes when and only when $a = \bar{d}$, $b = -\bar{c}$, that is, for a rotation of Σ_0 about an axis.

The transformation S carries I_S into a circle I'_S of the same size. I'_S is obviously the isometric circle of the inverse transformation S^{-1}.

The surface of the smaller polar cap of Σ_0 cut off by I_S we call the *interior of* I_S; that of the larger polar cap we call the *exterior of* I_S.

THEOREM 2. *Lengths and areas in the interior of I_S are increased in magnitude by the transformation S; those on the exterior of I_S are decreased.*

The caps cut off by I_S are carried into the equal caps cut off by I'_S. Since one cap is magnified and the other diminished, the smaller cap must undergo magnification and the larger diminution, whence the theorem. We have also

THEOREM 3. *The transformation S carries the interior and the exterior of I_S into the exterior and interior respectively of I'_S.*

Relative position of I_S and I'_S. If the linear transformation T corresponding to S is non-loxodromic, we can state the following facts about the position of I_S with respect to I'_S:

THEOREM 4. *If T is elliptic, I_S and I'_S intersect in the fixed points of S; if T is hyperbolic, I_S and I'_S are external; if T is parabolic, I_S and I'_S are tangent at the fixed point of S.*

This arrangement of the isometric circles on Σ_0 can be shown to exist by considering the relation between the fixed points and the fixed circles of the respective types of linear transformations and their projections on Σ_0.

Space transformations. We turn now to the space transformations resulting from successive inversions in an even number of spheres (or planes) orthogonal to Σ_0.

If two such transformations transform three points on Σ_0 in the same way, they transform all points of space in the same way.

If two sequences S and S' transform three points of Σ_0 alike, they transform all points of Σ_0 in the same way. For, when Σ_0 is projected stereographically on the plane there correspond to S and S' two linear transformations which transform three points alike; the two linear transformations are thus identical.

A point P of space, outside Σ_0 say, may be determined by three non-coaxal spheres orthogonal to Σ_0 through P. These intersect Σ_0 in three circles which are carried into the same three circles by S and S'. P is carried by both S and S' into the unique point P' outside Σ_0 lying on the three spheres orthogonal to Σ_0 through the latter three circles.*

Furthermore, there is a space transformation which produces any prescribed one-to-one and directly conformal transformation of Σ_0 on itself.

Consider S and its inverse S^{-1}. If Σ_0 is not rotated, S carries I_S into I'_S. Let Σ_S and Σ'_S be the spheres through these circles respectively, orthogonal to Σ_0. S must also transform Σ_S into Σ'_S of the *same size*.

---

* Poincaré, *loc. cit.* 4.

The sphere Σ_S is called the *isometric sphere* of the transformation S. Σ'_S is obviously the isometric sphere of S^{-1}.

Sequences Equivalent to S. We propose next to replace the sequence of inversions defining S by an equivalent system consisting, in general, of an inversion and either one or three reflections.

If S is a rotation it is equivalent to successive reflections in two diametral planes of Σ_0. This case we shall not consider.

Let P be a point on I_S and let P' be the point on I'_S into which S carries P. Since lengths on I_S are unaltered, if I_S be placed on I'_S so that P falls on P' and the orientation is correct, each point on I_S will fall on its corresponding point on I'_S.

As a point P moves counter-clockwise around I_S the corresponding point moves clockwise around I'_S. For, the interior of I_S goes into the exterior of I'_S, and the corresponding point must trace the boundary of the latter region in the same sense as the former is traced. Hence I_S must be turned over before being applied to I'_S.

Any sequence of an even number of inversions in spheres orthogonal to Σ_0, or reflections in diametral planes, is equivalent to S if it carries I_S into I'_S without altering lengths so that P falls on P' and so that the order of points about I'_S is opposite to that of the original points about I_S.

Let Π_S be the plane perpendicular bisector of the line of centers of Σ_S and Σ'_S, if these spheres are not coincident; if they coincide, let Π_S be the plane perpendicular bisector of the line joining any point P on I_S with its corresponding point P' on I'_S.

First, make an inversion in Σ_S. This leaves points on I_S unchanged. Second, make a reflection in Π_S, which reverses the direction of arcs on I_S. Finally, make a rotation about the line of centers of Σ'_S and Σ_0, so that the transform of P on I_S by the reflection in Π_S is brought into coincidence with P'. Then all the points of I_S are carried into their proper positions on I'_S. Since the rotation is equivalent to two reflections in planes, the result of the three operations is to leave angles on Σ_0 unchanged and to transform points on I_S into points on I'_S exactly as S transforms them.

These results may be stated as

THEOREM 5. *The space transformation S, if not a rotation, is equivalent to*

(1) *an inversion in Σ_S, followed by*

(2) *a reflection in Π_S, followed by*

(3) *a rotation through a suitable angle Θ about the fixed axis $0P'_S$ through the centers of Σ'_S and Σ_0.*

Let U_Π denote the reflection in Π_S; U_Σ, $U_{\Sigma'}$ the inversions in Σ_S, Σ'_S; and U_R, $U_{R'}$ the rotations of Σ_S, Σ'_S respectively, the directions of rotation being opposite. Then the transformations $U_{R'}U_\Pi U_{\Sigma'}$, $U_\Pi U_R U_{\Sigma'}$, $U_{\Sigma'}U_\Pi U_{R'}$, $U_\Pi U_\Sigma U_{R'}$, $U_{R'}U_{\Sigma'}U_\Pi$, and $U_{\Sigma'}U_{R'}U_\Pi$ are all equivalent since they transform three points on Σ_0 in the same way. Hence the sequence of the above theorem may be varied for a given transformation.

Remark. The center of an isometric sphere Σ_S is the transform of ∞ by S^{-1}, the inverse of S. For by the operations of Theorem 5 in the reverse order, ∞ remains fixed until the inversion (1) is performed, when it is carried to the center of Σ_S.

THEOREM 6. *Lengths inside Σ_S are increased in magnitude by the transformation S, those outside Σ_S are decreased.*

Of the geometric operations of Theorem 5 that accomplish S only the inversion in Σ_S changes lengths. Lengths within the sphere of inversion are increased, those without are decreased.

Remark. Any given sphere orthogonal to Σ_0 is the isometric sphere of an infinite number of possible space transformations; for example, the transformation resulting from an inversion in the given sphere, followed by a reflection in any diametral plane of Σ_0.

THEOREM 7. *The distance of a point P from the center of a sphere Σ_1 is unaltered by inversion in a sphere Σ if P lies on the surface of Σ_1 or on that of Σ; it is decreased if P lies outside both Σ_1 and Σ or inside both; it is increased if P lies outside one and inside the other.*

Without loss of generality we can take Σ as the unit sphere with center at the origin and Σ_1 with center $Q(q,0,0)$ on the ξ-axis. The spheres are given by

$$\Sigma:\ \xi^2+\eta^2+\zeta^2=1; \qquad \Sigma':\ (\xi-q)^2+\eta^2+\zeta^2=q^2-1.$$

Let $P(\xi,\eta,\zeta)$ be a point in space and $P'(\xi',\eta',\zeta')$ its inverse with respect to Σ. If h, h' denote the distances PQ, $P'Q$ respectively, then

$$h^2=(\xi-q)^2+\eta^2+\zeta^2, \qquad h'^2=(\xi'-q)^2+\eta'^2+\zeta'^2.$$

With the relation $\xi'=\xi/(\xi^2+\eta^2+\zeta^2)$ and similar ones for η', ζ', the difference $h^2-h'^2$ can easily be put into the form

$$h^2-h'^2=\frac{(\xi^2+\eta^2+\zeta^2-1)[(\xi-q)^2+\eta^2+\zeta^2-q^2+1]}{(\xi^2+\eta^2+\zeta^2)}.$$

The first factor in the numerator is positive for points outside Σ, negative for points inside, and zero for points on Σ. The same is true of the second factor for points with respect to Σ_1. The conclusions of the theorem are then evident.

Now consider the effect of a transformation S on the distance of a point P of space from the center of Σ_0. Of the three geometric operations of S only the inversion in Σ_S can affect this distance. Theorem 7 applied to the orthogonal spheres Σ_S and Σ_0 gives the following

THEOREM 8. *The distance of a point P from the center of Σ_0 is unaltered by S if P lies on either Σ_S or Σ_0; it is decreased if P lies outside or inside both; it is increased if P lies outside one and inside the other.*

Groups of space transformations. A group that contains infinitesimal transformations on Σ_0 is said to be *continuous on* Σ_0; that is, if given $\epsilon > 0$, there exists a transformation S of the group such that the distance between any point P on Σ_0 and its transform P' by S is less than ϵ. Otherwise the group is said to be *discontinuous on* Σ_0.

If for a discontinuous group there exists some point P on Σ_0 in whose neighborhood there is none of its transforms by the group, the group is *properly discontinuous on* Σ_0; if no such point exists, it is *improperly discontinuous on* Σ_0.

A study of groups of space transformations yields first two fundamental properties that characterize the continuous and the discontinuous groups on Σ_0.

THEOREM 9. *If a group contains an infinite number of rotations of Σ_0 into itself, it is continuous on Σ_0.*

Briefly, we can select a sequence of rotations $T_1, T_2, \cdots$, such that the axis of T_n approaches a limiting position and the angle of rotation θ_n approaches a limit. Then the rotation $T_{n+1}^{-1}T_n$, for n sufficiently large, alters the position of every point on Σ_0 by less than a preassigned small amount.

COROLLARY. *A group discontinuous on Σ_0 contains at most a finite number of rotations.*

THEOREM 10. *If for a group the number of isometric spheres of radii exceeding any positive number is infinite, the group is continuous on Σ_0.*

If the radii are bounded, we may select a sequence of transformations $T_1, T_2, \cdots$, where Σ_n approaches a limiting sphere, the sequence Σ'_n of the inverse transformations approaches a limiting sphere, and the angle of rotation of Theorem 5 approaches a limit. Then $T_{n+1}^{-1}T_n$, for n sufficiently large,

changes the position of every point on Σ_0 by less than a preassigned small amount.

If the radii are unbounded, the proof is similar, the limiting positions of Σ_n and Σ'_n being planes.

THEOREM 11. *For a group discontinuous on Σ_0 the number of isometric spheres of radii exceeding any positive number is finite.*

Transformations of a group. It can be shown by a proper choice of a transformation that a group may be transformed so that the resulting group contains no rotations of Σ_0 into itself.

In the following work it is assumed that such a transformation has been made on the group considered.

Limit points of a discontinuous group. Let S_k be the members of the group G discontinuous on Σ_0. Let r_k be the radius of Σ_{S_k}. Then by Theorem 11 any infinite sequence $r_1, r_2, \cdots$ of radii of distinct spheres has $\lim_{i=\infty} r_i = 0$.

A *limit point* of a group is defined to be a cluster point of centers of isometric spheres. Any point not a limit point of the group is called an *ordinary point*.

THEOREM 12. *The limit points of a discontinuous group lie on Σ_0.*

This is immediate, for the center of an isometric sphere whose radius approaches zero must approach the surface of Σ_0 since the spheres are orthogonal to Σ_0.

We say two figures F_1 and F_2 are *congruent with respect to a group* if $T(F_1) = F_2$, where $T \neq 1$ is a transformation of the group.

The region R of the group. A region R_0 is said to be a *fundamental region* of a discontinuous group if

(1) No two points of the region are congruent.

(2) In the neighborhood of any point on the boundary there are points congruent to points of the region.

The *region R* of the group is the space exterior to all isometric spheres of the group. A point P belongs to R if a region about P exists no point of which is interior to an isometric sphere of the group.

THEOREM 13. *No two points of R are congruent by any transformation of the group.*

For any point of R is carried by any member of the group into a point interior to the isometric sphere of the inverse transformation and so outside R.

THEOREM 14. *In the neighborhood of any limit point P of a discontinuous group lie an infinite number of distinct points congruent to any point of space with the possible exception of P and one other limit point.*

The proof of this theorem we omit.

THEOREM 15. *R and the regions congruent to R form a set of regions which extend into the neighborhood of every point of space.*

If this theorem is not true, there exists a closed sphere σ of ordinary points in which is no point of R or point congruent to R.* The points of σ are contained in a finite number of isometric spheres. Let σ' be a subsphere of σ wholly within the isometric spheres of $T_1, \cdots, T_n$ and no others. Consider $\sigma'_1 = T_1(\sigma)$. If a point of σ'_1 is within the isometric sphere of S then ST_1 increases lengths at a point of σ, so $ST_1 = T_k$. It follows that the points of σ' are contained in $n - 1$ isometric spheres at most; namely, those of the transformations $T_2T_1^{-1}, \cdots, T_nT_1^{-1}$. Making S we find a sphere congruent to a subsphere of σ', and hence congruent to a subsphere of σ, lying in $n - 2$ isometric spheres at most. Continuing in this way we arrive at a sphere congruent to a subsphere of σ which lies in no isometric spheres and hence is in R. This contradiction proves the theorem.

THEOREM 16. *R constitutes a fundamental region of the group.*

By Theorem 13 R satisfies the first condition for a fundamental region. In the neighborhood of any point P on the boundary of R there is a point Q within an isometric sphere in whose neighborhood by Theorem 15 are transforms of R. Thus R satisfies the second condition also.

THEOREM 17. *Any closed region not containing limit points of the group can be filled by a finite number of transforms of R, including possibly R itself. These regions fit together without lacunary spaces.*

Let B be the closed region. Since B contains no limit points, the number of isometric spheres containing points of B is finite, for an infinite sequence of spheres would have a cluster point of centers which would necessarily lie in B.

Each point of R is carried by a transformation T of the group into a point of R_T in the interior of Σ'_T. If Σ'_T contains no point of B, then R_T contains no point of B. Thus the number of regions R_T containing points of B is finite.

By Theorem 15 there can exist no lacunary spaces not filled by transforms of R.

* This method of proof is due to Professor L. R. Ford.

THEOREM 18. *Within any region enclosing a limit point there lie an infinite number of distinct transforms of the entire region R.*

Let P be the limit point and σ be a small sphere with P as center. Let Σ'_T be an isometric sphere lying entirely inside σ. Then $R_T = T(R)$ lies entirely within Σ'_T. Since σ contains an infinite number of isometric spheres, it contains an infinite number of transforms of R.

Groups with more than two limit points. If a group contains more than two limit points, the set of limit points is perfect. Let G be a group discontinuous on Σ_0 with more than two limit points. These points lie on Σ_0. Let P be a point on Σ_0 and U_P a spherical cap enclosing P. If every U_P contains limit points of G, for each P of Σ_0, the set of limit points is everywhere dense on Σ_0.

By Theorem 14 in the neighborhood of P lie an infinite number of transforms of P with two possible exceptions, P and one other limit point. P is exceptional only in that its transforms are not distinct. Thus in U_P there are transforms of P. If the set of limit points is everywhere dense on Σ_0, each point of Σ_0 has in its neighborhood transforms of itself. The group is therefore *improperly discontinuous on* Σ_0.

If the set of limit points is not everywhere dense on Σ_0, there is a closed cap U on Σ_0 consisting of ordinary points. Any point P of U which is not one of the finite number of possible fixed points has no transforms within a suitably small neighborhood of P. Hence a discontinuous group whose set of limit points is not everywhere dense on Σ_0 is *properly discontinuous on* Σ_0.

THEOREM 19. *A necessary and sufficient condition that a group be properly discontinuous on Σ_0 is that the region R contain in its interior a point of Σ_0.*

A point P of Σ_0 within R has no transforms within some cap U_P about P. Conversely, if the group is properly discontinuous on Σ_0, there exists at least one point P on Σ_0 and a sphere σ about P which contains only ordinary points. The sphere σ can be filled by a finite number of transforms of R. Let R_m be one of them. R_m has a finite number of bounding spherical surfaces in σ, each orthogonal to Σ_0. At least one point P' on Σ_0 exists interior to R_m. Hence $T_m^{-1}(P') = P_1$, where P_1 is interior to R and lies on Σ_0.

Thus for a group improperly discontinuous on Σ_0 no point of the surface of Σ_0 belongs to the interior of R. Hence R is divided into two regions, one inside, the other outside Σ_0, having at most common boundary points on Σ_0.

The boundary of R. A point P of the boundary of R is either a limit point or an ordinary point. An ordinary boundary point lies on a finite number of isometric spheres and is inside none.

We consider below the properties of ordinary boundary points of R; those of the limit points have been discussed above.

An ordinary boundary point P belongs to one of the following classes:

(α) P on just one isometric sphere.

(β) P on two or more isometric spheres having a circle in common.

(γ) P on a finite number of isometric spheres not having a circle in common.

(δ) P a point of tangency of two or more isometric spheres and on no isometric spheres not tangent at P.

If P is a fixed point of $S = S^{-1}$ and on no other isometric sphere than Σ_S, it shall be regarded as belonging to class (β).

Points of Class (α). Let P lie on the isometric sphere Σ_S of S. Then $P' = S(P)$ lies on Σ'_S and is distinct from P.

If P' is inside an isometric sphere Σ_U of U, then US increases lengths near P', for S does not alter lengths near P and U increases lengths near P'. Hence P lies inside Σ_{US}, which is contrary to hypothesis. P' is therefore a boundary point of R.

P' belongs to class (α) also. For, if P' were on Σ_U, $U \neq S^{-1}$, then US would leave lengths near P unchanged since S does not alter lengths near P; nor U those near P'. Thus P would be on Σ_{US} also, which is contrary to the fact.

Points on Σ_S in the neighborhood of P are ordinary boundary points. Part of the boundary of R is therefore a portion of the surface of Σ_S and a congruent portion of Σ'_S. The two congruent parts are equal in area since lengths, and consequently areas, on Σ_S are unchanged by S. This part of Σ_S limited by boundary points of other classes is called a *face*.

These results with Theorem 8 give

THEOREM 20. *The boundary points of R of class (α) form sets of spherical faces which are congruent in pairs. The congruent faces are equal in area and congruent points on the faces are equidistant from the center of Σ_0.*

Points of Class (β). Let P lie on C_1, the circle common to the two or more isometric spheres. Let E_1 be the arc (or arcs) of C_1 which consists entirely of ordinary boundary points of R. We call E_1 an *edge* of R.

Let $T_1, T_2, \cdots, T_m$ be the transformation whose isometric spheres pass through E_1. Then $E_k = T_k(E_1)$, $k = 2, \cdots, m$, is an edge congruent to E_1 on the boundary of R, and there are no other congruent edges. That E_k is on the boundary of R follows as in class (α). From the way lengths on E_k are affected we see that T_k^{-1}, $T_1T_k^{-1}, \cdots, T_{k-1}T_k^{-1}$, $T_{k+1}T_k^{-1}, \cdots, T_mT_k^{-1}$ have

isometric spheres passing through E_k, and that E_k is external to all other isometric spheres.

Since T_k carries E_1 into E_k with no change in length we have

THEOREM 21. *The boundary points of class* (β) *form sets of congruent circular arcs, or edges. All congruent edges are equal in length.*

Points of Class (γ). P of class (γ) is called a *vertex.* Let P_1 be on the isometric sphere Σ_{T_1} which forms a face F_1 of R. Let $T_1(P_1) = P_2$. Since $T_1(F_1) = F'_1$ of the same area and shape, that is, a spherical polygon on Σ'_{T_1}, at least two other isometric spheres must pass through P_2, each forming a face of R. Hence P_2 is also a vertex.

Let $T_1, T_2, \cdots, T_m$ be the transformations whose isometric spheres form the faces at P_1. Then $P_k = T_k(P_1)$ is a vertex. Conversely, to each vertex congruent to P_1 by U say, there corresponds an isometric sphere Σ_U through P_1. For, P_1 on the boundary of R cannot lie inside Σ_U; if it lies outside Σ_U then $P' = U(P_1)$ lies inside Σ'_U and not on the boundary of R. The number of vertices congruent to P_1 is therefore finite.

Applying also Theorem 8 we have

THEOREM 22. *The boundary points of class* (γ) *form finite sets of congruent vertices. All vertices of a set are equidistant from the center of* Σ_0.

Points of Class (δ). A point of tangency of two isometric spheres lies necessarily on Σ_0. One or two of the tangent spheres form faces of R according as the spheres lie on the same or opposite sides of the tangent plane. Boundary points of this class do not occur for a group improperly discontinuous on Σ_0.

Angles at sets of congruent edges and vertices.

THEOREM 23. *The sum of the dihedral angles at the edges of a set of congruent edges is* 2π *or a sub-multiple of* 2π.

Let $E_1, E_2, \cdots, E_m$ be the set of edges congruent respectively by $T_1, T_2, \cdots, T_m$, with $T_m(E_m) = E_1$. Let the faces at E_1 be F_0, F_1.

The transformation $U_1 = T_m T_{m-1} \cdots T_2 T_1$ has E_1 as fixed edge. Let

$$U_m = T_m;\quad U_{m-1} = T_m T_{m-1};\quad \cdots;\quad U_2 = T_m T_{m-1} \cdots T_2;$$

and

$$U_1 = T_m T_{m-1} \cdots T_2 T_1.$$

Of these U_m carries E_m into E_1, the dihedral angle in R at E_m into an equal angle at E_1, and the region R into R_m joining onto R along the face F_0. U_{m-1} carries E_{m-1} into E_1, the dihedral angle about E_{m-1} into an equal one at E_1, and R into R_{m-1}, fitting onto R_m along the open face of R_m through E_1.

Finally, U_1 carries E_1 into E_1, the dihedral angle at E_1 into an equal one at E_1, and R into R_1, fitting onto R_2 along its open face through E_1.

If R_1 coincides with R, then U_1 is the identical transformation and the sum of the dihedral angles at the edges is 2π.

If $U_1 \neq 1$, we make $U_1, U_1^2, U_1^3, \cdots$ Each of these has an isometric sphere through E_1, so the number of possible repetitions of U_1 is finite. Then with some U_1^k the dihedral angle about E_1 is completely filled. The angles of the cycle have been used $k+1$ times; hence their sum is $2\pi/(k+1)$.

THEOREM 24. *The sum of the solid angles at the vertices of a set of congruent vertices is $4\pi/m$, m a positive integer.*

It is a property of conformal space transformations that solid angles are preserved.

Make the inverses of all the transformations with isometric spheres through P_1. R and the transforms of points of R near the vertices of a cycle fill out the solid angle about P_1. Each solid angle at a vertex has the same number of transforms at P_1. Thus if $T_1, \cdots, T_m$ carry P_k to P_1 and if S carries P_l to P_k, then $T_1S, \cdots, T_mS$ and no others carry P_l to P_1. Then the sum of the solid angles of the set is $4\pi/m$.

Generating transformations of the group. A set of transformations $T_1, T_2, \cdots$ is said to *generate* a group G if every transformation of G is a combination of the set.

The following theorem can be readily proved:

THEOREM 25. *The set of transformations which connect the faces of R form a set of generating transformations of the group.*

Applications to properly discontinuous groups in the plane. Let M be a properly discontinuous group of linear transformations. A sufficient condition for finding a generating set for M is that it be possible to join every ordinary point to a point of R by a curve not passing through a limit point. In such a case, the transformations connecting the sides of R form a generating set for M.

If this condition is not satisfied, we are able to find a generating set for M as follows. Project the plane stereographically on Σ_0 and form the corresponding space group G. The region R' for G is found by constructing the isometric spheres. R' must contain in its interior a point of Σ_0, for otherwise G, and hence M, would be improperly discontinuous. A generating set of transformations for G is the set connecting the faces of R'. By projecting Σ_0 back on the plane those parts of Σ_0 belonging to R' form a region R for M (possibly disconnected). The set of linear transformations corresponding to the generating set for G forms a generating set for M.

ON SUMMABILITY OF DOUBLE SEQUENCES.*

By Ralph Palmer Agnew.†

1. *Introduction.* Let $\| a_{mi} \|$ and $\| b_{nj} \|$ be two triangular matrices of real or complex constants satisfying the conditions

(1.1) for each i, $\lim_{m \to \infty} a_{mi} = 0$; for each j, $\lim_{n \to \infty} b_{nj} = 0$;

(1.2) for each m, $\sum_{i=0}^{m} | a_{mi} | < K$; for each n, $\sum_{j=0}^{n} | b_{nj} | < K$,

K being a constant independent of m and n, and

(1.3) $\lim_{m,n \to \infty} C_{mn} = 1$ where $C_{mn} = \sum_{i=0}^{m} \sum_{j=0}^{n} a_{mi} b_{nj}$.

With each convergent or divergent double sequence s_{ij}, we associate a transform S_{mn} defined by

$$F \equiv F(a, b) \qquad S_{mn} = \sum_{i=0}^{m} \sum_{j=0}^{n} a_{mi} b_{nj} s_{ij}.$$

A sequence s_{ij} is said to be "summable F" to S if its transform S_{mn} converges to S, to be "bounded F" if S_{mn} is uniformly bounded for all m and n, and to be "ultimately bounded F" if $\limsup_{m,n \to \infty} | S_{mn} | < \infty$.

Results of G. M. Robison ‡ show that each *bounded* convergent sequence must be summable F to the value to which it converges; but results of T. Kojima § show that conditions much more severe than (1.1), (1.2) and (1.3) are necessary to ensure that each convergent sequence shall be summable F. Hence F may be non-regular, for the simple reason that there may be unbounded convergent sequences which it fails to evaluate.

The transformation F is of the type called *factorable* by C. R. Adams, and has been investigated by C. R. Adams ¶ and F. Lösch.‖ An important special case has been considred by S. Bochner.**

* Presented to the American Mathematical Society, March 25, 1932.

† National Research Fellow.

‡ G. M. Robison, *Transactions of the American Mathematical Society*, Vol. 28 (1926), pp. 50-73.

§ T. Kojima, *Tohoku Mathematical Journal*, Vol. 21 (1922), pp. 3-14.

¶ C. R. Adams, I, *Bulletin of the American Mathematical Society*, Vol. 37 (1931), pp. 741-748; II, *Transactions of the American Mathematical Society*, Vol. 34 (1932), pp. 215-230.

‖ F. Lösch, *Mathematische Zeitschrift*, Vol. 34 (1931), pp. 281-290. Lösch postulates, instead of (1.3), the two conditions $\lim_{m \to \infty} \sum_{i=0}^{m} a_{mi} = 1$, $\lim_{n \to \infty} \sum_{j=0}^{n} a_{nj} = 1$; this difference is, however, as Adams II points out, quite trivial.

** S. Bochner, *Mathematische Zeitschrift*, Vol. 35 (1932), pp. 122-126.

Lösch (*loc. cit.*, p. 282) has shown that *if s_{ij} converges to s and S_{mn} is bounded, then s_{ij} is summable F to s;* and (p. 287) that *if s_{ij} converges to s and is summable F to S, then $S = s$.* These results are of great interest since on one hand they show an extent to which unbounded convergent sequences are summable F, and on the other hand that F is consistent with convergence. The latter result has important applications in the theory of series of functions; for example, if a double trigonometric series is summable F to a function $f(x, y)$, then the series must converge to $f(x, y)$ for all values of x and y for which it converges.

It is the object of the present paper to prove and discuss the following and a related theorem.

THEOREM 1. *If s_{ij} converges to s and if there exist an index Q and two sequences α_m and β_n of constants such that*

$$\text{for each } m > Q, \qquad | S_{mn} | < \alpha_m, \qquad n > Q, \tag{1.4}$$

$$\text{for each } n > Q, \qquad | S_{mn} | < \beta_n, \qquad m > Q, \tag{1.5}$$

then s_{ij} is summable F to s.

This theorem may be stated as follows. *If a double sequence converges and if each sufficiently advanced row and column of its F-transform is bounded, then the sequence is summable F to the value to which it converges.*

2. *Consequences of Theorem* 1. Before passing to a proof of Theorem 1, we give two of its corollaries in Theorems 2 and 3.

THEOREM 2. *If s_{ij} converges to s and is ultimately bounded F, then s_{ij} is summable F to s.**

THEOREM 3. *If s_{ij} converges to s and is summable F to S, then $S = s$.*

Theorem 2 contains the first result of Lösch mentioned in § 1; Theorem 3 is, except for the fact that our transformations are slightly more general than those of Lösch, precisely the second result of Lösch.

Theorem 1 also contains the following theorem of Adams I, p. 743; if s_{ij} converges to s, and if there exist sequences $A_0, A_1, A_2, \cdots$ and $B_0, B_1, B_2, \cdots$ of constants such that for each $j \geqq 0$ and each $i \geqq 0$ and for all m and n,

$$| \sum_{i=0}^{m} a_{mi} s_{ij} | < A_j; \qquad | \sum_{j=0}^{n} b_{nj} s_{ij} | < B_i, \tag{2.1}$$

---

* It follows at once from (1.2) that each bounded sequence is bounded F; hence Theorem 2 includes the result of Robison that each bounded convergent sequence is summable F to the value to which it converges.

then S_{mn} converges to S. For, suppose the hypotheses of Adams' theorem hold; then for each fixed $m \geqq 0$ we have

$$(2.2) \qquad | S_{mn} | \leqq \sum_{i=0}^{m} | a_{mi} | \ | \sum_{j=0}^{n} b_{nj} s_{ij} | \leqq \sum_{i=0}^{m} | a_{mi} | B_i.$$

An analogous set of inequalities holds for each fixed $n \geqq 0$. Thus we see that the hypotheses of Theorem 1 hold with $Q = -1$.*

3. *A Lemma.* The following lemma will be used in proofs of our theorems.

LEMMA 1. *Let R be a non-negative integer, let g_{mi} and G_{in} be double sequences of real or complex constants, and let*

$$(3.1) \qquad \lim_{m \to \infty} g_{mi} = 0, \qquad (i = 0, 1, 2, \cdots).$$

If there is an index N and a sequence H_m of constants such that whenever $n > N$ we have

$$(3.2) \qquad | g_{m0} G_{0n} + g_{m1} G_{1n} + \cdots + g_{mR} G_{Rn} | < H_m, \qquad m > N,$$

then

$$(3.3) \qquad \lim_{m,n \to \infty} (g_{m0} G_{0n} + g_{m1} G_{1n} + \cdots + g_{mR} G_{Rn}) = 0.$$

We prove this lemma by induction. It is easy to show, by considering separately the case where $g_{m0} = 0$ for all $m > N$ and the case where $g_{\mu 0} \neq 0$ for some fixed $\mu > N$, that the lemma holds when $R = 0$.

Assuming that the lemma holds when $R = 0, 1, 2, \cdots, \rho - 1$, we prove that it holds for $R = \rho$ by considering the infinite matrix

$$(3.4) \qquad \| g_{mi} \| \qquad (m = N + 1, n + 2, \cdots; \ i = 0, 1, 2, \cdots, \rho).$$

If this matrix has rank less than $(\rho + 1)$, then the columns are linearly

---

* We give a few remarks bearing on the relations between our Theorem 1 and Theorem 2 of Adams II. A transformation T is said to be regular for a class $\mathcal{S}$ of sequences if each convergent sequence belonging to $\mathcal{S}$ is summable T to the value to which it converges. With this terminology, Adams shows that F is regular for the class $\mathcal{B}$ of all sequences which are bounded F. Now our Theorem 1 shows that F is regular for the class $\mathcal{L}$ of all sequences having F-transforms of which each sufficiently advanced row and column is bounded.

It is clear that $\mathcal{L}$ contains all sequences which are summable F; hence $\mathcal{L}$ is, apart from divergent sequences, the largest class of sequences for which F is regular. It is also clear that in case F has an inverse, and in many other cases as well, a convergent sequence may belong to $\mathcal{L}$ and fail to belong to $\mathcal{B}$. Hence Theorem 2 of Adams II, and the later Theorems of Adams II which depend upon it, can be made stronger as well as easier to apply by using the class $\mathcal{L}$ instead of the class $\mathcal{B}$.

dependent, i. e. there exists a system $\lambda_0, \lambda_1, \cdots, \lambda_\rho$ of constants (not all zero) such that

$$\lambda_0 g_{m0} + \lambda_1 g_{m1} + \cdots + \lambda_\rho g_{m\rho} = 0, \qquad m > N.$$

Selecting an index α such that $\lambda_\alpha \neq 0$, we obtain

$$g_{m\alpha} = (\lambda_0 g_{m0} + \cdots + \lambda_{\alpha-1} g_{m,\alpha-1} + \lambda_{\alpha+1} g_{m,\alpha+1} + \cdots + \lambda_R g_{mR})/\lambda_\alpha.$$

When we substitute for $g_{m\alpha}$ in the relations obtained by setting $R = \rho$ in (3. 2) and (3. 3), we find that we have reduced our problem to the case $R < \rho$.

If on the other hand the matrix (3. 4) has rank $(\rho + 1)$, let

$$(3.5) \qquad \det(g_{m_\alpha i}) \qquad (\alpha = 0, 1, \cdots, \rho;\ i = 0, 1, \cdots, \rho)$$

be a non-vanishing $(\rho+1)$-rowed determinant selected from its elements. Let us now consider only values of n which exceed N. Since $m_0, m_1, \cdots, m_\rho$ all exceed N, we have by (3. 2)

$$(3.6) \qquad | g_{m_\alpha 0} G_{0n} + g_{m_\alpha 1} G_{1n} + \cdots + g_{m_\alpha \rho} G_{\rho n} | < H_{m_\alpha}, \quad (\alpha = 0, 1, \cdots, \rho).$$

Let us set

$$(3.61) \qquad g_{m_\alpha 0} G_{0n} + g_{m_\alpha 1} G_{1n} + \cdots + g_{m_\alpha \rho} G_{\rho n} = H_{m_\alpha n} \quad (\alpha = 0, 1, \cdots, \rho).$$

Then by (3. 6) we have for all values of n under consideration

$$(3.7) \qquad | H_{m_\alpha n} | < H_{m_\alpha}, \qquad (\alpha = 0, 1, \cdots, \rho).$$

Since (3. 5) does not vanish, we can solve the equations (3. 61) for G_{in} obtaining

$$(3.8) \qquad G_{in} = \Delta_{i0} H_{m_0 n} + \Delta_{i1} H_{m_1 n} + \cdots + \Delta_{i\rho} H_{m_\rho n}, \quad (i = 0, 1, \cdots, \rho).$$

where the Δ_{ij} are constants depending only on the elements of the determinant (3. 5). It follows from (3. 7) and (3. 8) that each of the sequences

$$(3.9) \qquad G_{0n}, G_{1n}, \cdots, G_{\rho n}$$

is bounded for all $n > N$; hence we may use (3. 1) to obtain (3. 3) for $R = \rho$, and the proof by induction is complete.

4. *Proof of Theorem* 1. Let s_{ij} be any given sequence converging to s and having an F-transform satisfying the hypotheses of Theorem 1. Given $\epsilon > 0$, choose an index R which is greater than Q and also so great that

$$(4.1) \qquad | s_{mn} - s | < \epsilon/2K^2, \qquad m, n > R,$$

K being the constant in (1. 2). When $m, n > R$, we have

$$S_{mn} - s = \{\sum_{i=0}^{R} \sum_{j=0}^{n} + \sum_{i=0}^{m} \sum_{j=0}^{R} - \sum_{i=0}^{R} \sum_{j=0}^{R} + \sum_{i=R+1}^{m} \sum_{j=R+1}^{n}\} a_{mi} b_{nj} (s_{ij} - s) + s(C_{mn} - 1)$$
$$= S_{mn}^{(1)} + S_{mn}^{(2)} - S_{mn}^{(3)} + S_{mn}^{(4)} + s(C_{mn} - 1).$$

Using (4.1) and (1.2), we see that $| S_{mn}^{(4)} | < \epsilon/2$. Also (1.1) and (1.3) imply that we can choose $r > R$ so great that $| -S_{mn}^{(3)} + s(C_{mn} - 1) | < \epsilon/2$ when $m, n > r$. Then on one hand

$$(4.2) \qquad | S_{mn} - s | < | S_{mn}^{(1)} | + | S_{mn}^{(2)} | + \epsilon, \qquad m, n > r$$

and on the other hand

$$(4.3) \qquad | S_{mn}^{(1)} + S_{mn}^{(2)} | < | S_{mn} | + | s | + \epsilon, \qquad m, n > r.$$

We proceed to show that

$$(4.4) \qquad \lim_{m,n\to\infty} S_{mn}^{(1)} = 0.$$

Using (1.1), we see that for each fixed m, $\lim_{n\to\infty} S_{mn}^{(2)} = 0$; hence there is a sequence H'_m of constants such that for each m,

$$(4.5) \qquad | S_{mn}^{(2)} | < H'_m, \qquad n > r.$$

Combining (1.4), (4.3), and (4.5) we obtain for each $m > r$

$$(4.6) \qquad | S_{mn}^{(1)} | < H_m \qquad n > r$$

where

$$(4.7) \qquad H_m = H'_m + \alpha_m + | s | + \epsilon.$$

Introducing the notation

$$(4.8) \qquad A_{in} = \sum_{j=0}^{n} b_{nj}(s_{ij} - s),$$

in (4.6), we obtain for each fixed $m > r$

$$(4.9) \qquad | S_{mn}^{(1)} | = | a_{m0}A_{0n} + a_{m1}A_{1n} + \cdots + a_{mR}A_{Rn} | < H_m, \qquad n > r.$$

An application of Lemma 1 yields (4.4). An analogous argument shows that $\lim_{m,n\to\infty} S_{mn}^{(2)} = 0$. It therefore follows from (4.2) that $\lim_{m,n\to\infty} S_{mn} = s$ and Theorem 1 is proved.

5. *A variation of Theorem* 1. Lösch, *loc. cit.*, pp. 285-287, imposes upon the elements of $\| a_{mi} \|$ and $\| b_{nj} \|$ the following condition. Corresponding to each pair q and Q of indices, there exist two systems $m_0 < m_1 < \cdots < m_q$ and $n_0 < n_1 < \cdots < n_q$ of indices such that $m_0 > Q$, $n_0 > Q$, and each of the determinants

$$(5.1) \qquad \det(a_{m_j i}); \quad \det(b_{n_i j}) \qquad (i = 0, 1, \cdots, q;\ j = 0, 1, \cdots, q)$$

does not vanish. Lösch shows that when this condition, which we shall designate by (5.1), as well as (1.1), (1.2), and (1.3) hold, then each convergent sequence summable F must be also bounded F. We now give a theorem which includes this result.

THEOREM 4. *Let F satisfy (5.1) as well as (1.1), (1.2), and (1.3). If s_{ij} converges and has a transform satisfying the hypotheses of Theorem 1, then there exist an index r and two sequences α'_m and β'_n of constants such that*

$$\text{(5.2)} \qquad \textit{for each } m, \ |S_{mn}| < \alpha'_m, \qquad n > r,$$

and

$$\text{(5.3)} \qquad \textit{for each } n, \ |S_{mn}| < \beta'_n, \qquad m > r.$$

This theorem may be stated as follows. *Let F satisfy (5.1) as well as (1.1), (1.2), and (1.3). If s_{ij} converges and if each sufficiently advanced row and column of its F-transform is bounded, then each row and column of its F-transform is bounded.* It follows that, when F satisfies these conditions, *each convergent sequence which is ultimately bounded F is also bounded F.*

To prove Theorem 4, we proceed precisely as in the proof of Theorem 1 to obtain (4.9). The hypothesis (5.1) ensures the existence of a system $R < m_0 < m_1 < \cdots < m_R$ of indices such that the determinant

$$\text{(5.4)} \qquad \det(a_{m_j i}) \qquad (i = 0, 1, \cdots, R;\ j = 0, 1, \cdots, R)$$

does not vanish. Using the fact that (4.9) holds when $m = m_0, m_1, \cdots, m_R$, we see as in the latter part of the proof of Lemma 1 that each of the sequences

$$\text{(5.5)} \qquad A_{0n}, A_{1n}, \cdots, A_{Rn}$$

is bounded for all $n > r$. Since this is a finite set of sequences they are uniformly bounded for all $n > r$, i. e. there is a constant A_R such that whenever $i \leqq R$, we have

$$\text{(5.6)} \qquad |A_{in}| < A_R \qquad n > r.$$

Hence when $m \leqq R$ we have

$$\text{(5.7)} \qquad |S_{mn}| \leqq \sum_{i=0}^{m} |a_{mi}|\ |A_{in}| < KA_R \qquad n > r.$$

But R was chosen greater than Q, and r greater than R; hence it results from (1.4) that when $m > R$,

$$\text{(5.8)} \qquad |S_{mn}| < \alpha_m \qquad n > r.$$

Letting $\alpha'_m = KA_R$ when $m \leqq R$ and $\alpha'_m = \alpha_m$ when $m > R$, we see that (5.2) follows from (5.7) and (5.8). An analogous argument yields (5.3) and Theorem 4 is proved.

6. *Applications.* Let V and W represent any methods which associate with each double sequence a transformed double sequence, and suppose

$$\text{(6.1)} \qquad V = F(a, b) W$$

where F is defined as in § 1 and FW represents the transformation which associates with a sequence the F-transform of its W-transform. An application of Theorem 1 gives the following result. *If $V = FW$, and if s_{ij} is summable W to s and each sufficiently advanced row and column of its V-transform is bounded, then s_{ij} is summable V to s.* A corollary of this result gives an application of Theorem 2, namely, *if $V = FW$ and if s_{ij} is summable W to s and ultimately bounded V, then s_{ij} is summable V to s.* A further corollary gives an application of Theorem 3, namely, *if $V = FW$, and s_{ij} is summable W to s and summable V to S, then $S = s$;* in other words *V and W are consistent.*

There is an important class of transformations V and W for which (6.1) holds. Let

$$A^{(\alpha)}:\quad A_n^{(\alpha)} = \sum_{k=0}^{n} a_{nk}^{(\alpha)} s_k; \qquad B^{(\alpha)}:\quad B_n^{(\alpha)} = \sum_{k=0}^{n} b_{nk}^{(\alpha)} s_k \qquad (\alpha = 1, 2)$$

be four simple-sequence transformations about which nothing is assumed other than that $A^{(1)} = C^{(1)}B^{(1)}$ and $A^{(2)} = C^{(2)}B^{(2)}$ where $C^{(1)}$ and $C^{(2)}$ are regular transformations with triangular matrices $\| c_{nk}^{(1)} \|$ and $\| c_{nk}^{(2)} \|$. It is easy to show that

$$(6.2) \qquad A^{(1)} \odot A^{(2)} = (C^{(1)} \odot C^{(2)})\,(B^{(1)} \odot B^{(2)})$$

where $A^{(1)} \odot A^{(2)}$ is the double-sequence transformation defined by *

$$A_{mn} = \sum_{i=0}^{m} \sum_{j=0}^{n} a_{mi}^{(1)} a_{nj}^{(2)} s_{ij}$$

and $B^{(1)} \odot B^{(2)}$ and $C^{(1)} \odot C^{(2)}$ are similarly defined. The transformations $A^{(1)} \odot A^{(2)}$ and $B^{(1)} \odot B^{(2)}$ need not have the form F since the conditions analogous to (1.1), (1.2) and (1.3) may fail to hold. However regularity of $C^{(1)}$ and $C^{(2)}$ ensures that $C^{(1)} \odot C^{(2)}$ is of the form F; hence (6.2) is of the form (6.1) and our results may be applied. We state, for reference, the following theorems.

THEOREM 5. *Let $A^{(1)} = C^{(1)}B^{(1)}$ and $A^{(2)} = C^{(2)}B^{(2)}$ where $C^{(1)}$ and $C^{(2)}$ are regular. If s_{ij} is summable $B^{(1)} \odot B^{(2)}$ to s, and if each sufficiently advanced row and column of the $A^{(1)} \odot A^{(2)}$ transform of s_{ij} is bounded, then s_{ij} is summable $A^{(1)} \odot A^{(2)}$ to s.*

THEOREM 6. *If $A^{(1)} = C^{(1)}B^{(1)}$ and $A^{(2)} = C^{(2)}B^{(2)}$ where $C^{(1)}$ and $C^{(2)}$ are regular, then $A^{(1)} \odot A^{(2)}$ and $B^{(1)} \odot B^{(2)}$ are consistent.*

7. *The unsymmetric case.* Let I represent the identity transformation. Theorem 3 shows that if A and B are regular, then $A \odot B$ and $I \odot I$ are

* This is the notation of Adams I and II.

consistent. We now propose to prove and give consequences of the following theorem.

THEOREM 7. *If A and B are regular, then $A \odot I$* and *$I \odot B$ are consistent.*

The proof of Theorem 7 follows at once from Theorem 5 and the following lemma.

LEMMA 2. *Let A and B be transformations satisfying* (1.2). *If each sufficiently advanced row of the $A \odot I$ transform of s_{ij} is bounded and each sufficiently advanced column of the $I \odot B$ transform of s_{ij} is bounded, then each sufficiently advanced row and column of the $A \odot B$ transform of s_{ij} is bounded.*

To prove this lemma, let s_{ij} be a sequence satisfying its hypotheses, and choose sequences α_m and β_n of constants and an index Q such that for each $m > Q$

$$(7.1) \qquad \left| \sum_{i=0}^{m} a_{mi} s_{in} \right| < \alpha_m \qquad n > Q$$

and for each $n > Q$

$$(7.2) \qquad \left| \sum_{j=0}^{n} b_{nj} s_{mj} \right| < \beta_n \qquad m > Q.$$

Letting S_{mn} represent the $A \odot B$ transform of s_{ij}, we may write for each fixed $n > Q$

$$S_{mn} = \sum_{i=0}^{Q} \sum_{j=0}^{n} a_{mi} b_{nj} s_{ij} + \sum_{i=Q+1}^{m} a_{mi} \sum_{j=0}^{n} b_{nj} s_{ij}.$$

Using (1.2) and (7.2), we obtain

$$S_{mn} \leqq K^2 \max_{0 \leq i \leq Q;\ 0 \leq j \leq n} | s_{ij} | + K\beta_n \qquad m > Q.$$

The right member of this inequality depends only on n; hence each column of S_{mn} with a fixed index $n > Q$ is bounded. An analogous argument shows that each row of S_{mn} with a fixed index $m > Q$ is bounded and the lemma is proved.

A transformation A is said to include a transformation B (written $A \supset B$) if each sequence summable B is also summable A to the same value. From Theorem 7, we obtain the following more inclusive result.

THEOREM 8. *If $A^{(1)} \supset B^{(1)}$ and $B^{(2)} \supset A^{(2)}$, and $B^{(1)}$ and $A^{(2)}$ have inverses, then $A^{(1)} \odot A^{(2)}$ and $B^{(1)} \odot B^{(2)}$ are consistent.*

Since $A^{(1)}[B^{(1)}]^{-1}$ and $B^{(2)}[A^{(2)}]^{-1}$ are regular it follows from Theorem 7 that $A^{(1)}[B^{(1)}]^{-1} \odot I$ and $I \odot B^{(2)}[A^{(2)}]^{-1}$ are consistent; hence

$$A^{(1)} \odot A^{(2)} = \{A^{(1)}[B^{(1)}]^{-1} \odot I\}\{B^{(1)} \odot A^{(2)}\}$$

and

$$B^{(1)} \odot B^{(2)} = \{I \odot B^{(2)}[A^{(2)}]^{-1}\}\{B^{(1)} \odot A^{(2)}\}$$

are consistent and Theorem 8 is proved.

8. *Conclusion.* Combining Theorems 6 and 8 we obtain

THEOREM 9. *If $A^{(1)}$ includes or is included by $B^{(1)}$, if $A^{(2)}$ includes or is included by $B^{(2)}$, and if $A^{(1)}$, $B^{(1)}$, $A^{(2)}$, and $B^{(2)}$ all have inverses, then $A^{(1)} \odot A^{(2)}$ and $B^{(1)} \odot B^{(2)}$ are consistent.*

Theorems 5, 6, 8, and 9 have many immediate applications. We mention only a few of the applications of Theorem 9 to Cesàro and Holder methods of summability.

Let $C(r)$ and $H(r)$ denote respectively the simple-sequence Cesàro and Holder transformations of order r; then the double-sequence Cesàro and Holder transformations $C(r_1, r_2)$ and $H(r_1, r_2)$ are defined to be $C(r_1) \odot C(r_2)$ and $H(r_1) \odot H(r_2)$ respectively. Applying Theorem 9, we see that *if r_1, r_2, r_3, and r_4 are all real and greater than -1, then any two of the transformations $C(r_1, r_2)$, $C(r_3, r_4)$, $H(r_1, r_2)$, and $H(r_3, r_4)$ are consistent.** To this set of consistent methods can be added methods of the form $C(r_1) \odot H(r_2)$ and $H(r_1) \odot C(r_2)$.

BROWN UNIVERSITY,
CORNELL UNIVERSITY.

* This application of Theorem 9 depends upon the fact that when r and r' are real and greater than -1, we have at least one of the relations $C(r) \supset C(r')$ and $C(r') \supset C(r)$, and the fact that $C(r)$ and $H(r)$ are equivalent when $r > -1$. For references to literature, see E. Kogbetliantz, *Sommation des séries et intégrales divergentes par les moyennes arithmétiques et typiques*, Paris (1931), pp. 17-19. It should be noted that we cannot establish a part of our consistency theorem by showing that $C(r', \rho') \supset C(r, \rho)$ when $r' > r > -1$ and $\rho' > \rho > -1$; that the latter result does not hold follows from the fact that $C(1,1)$ and $C(0,0)$ are overlapping methods of summability in the sense that each evaluates certain sequences which the other fails to evaluate. It should be noted also that the equivalence theorem for simple-sequence Cesàro and Holder methods cannot be extended to double-sequence transformations. In fact we can show that $C(2,2)$ and $H(2,2)$ are overlapping methods of summability by using the two identities

$$C(2,2) = \{C(2)[H(2)]^{-1} \odot C(2)[H(2)]^{-1}\} H(2,2)$$

and

$$H(2,2) = \{H(2)[C(2)]^{-1} \odot H(2)[C(2)]^{-1}\} C(2,2)$$

and the Kojima (*loc. cit.*) conditions for regularity of double sequence transformations.

ON THE EXISTENCE OF CRITICAL POINTS OF GREEN'S FUNCTIONS FOR THREE-DIMENSIONAL REGIONS.

By Tsai-Han Kiang.

1. *Introduction.* The regions we shall consider here are the connected, closed, and bounded three-dimensional regions, for which the Dirichlet problem is possible and which are 3-complexes * in the sense of analysis situs. For simplicity let us call such a region an *admissible region.* Of an admissible region the connectivity numbers R_0 and R_3 are obviously 1 and 0 respectively. We shall investigate the critical points † of Green's functions first for general admissible regions with R_1 and R_2 not both zero (Theorem 1), then for admissible regions with R_1 and R_2 both zero but not homeomorphic with a spherical region (Theorem 2), and finally for admissible regions not only with R_1 and R_2 both zero but also homeomorphic with a spherical region (Theorem 4).‡

2. *Critical points of Green's functions for general admissible regions.* For our present purpose we shall state a special case of a theorem in a previous paper by the author § as the following theorem:

Theorem A. *Suppose a non-degenerate* ¶ *function* $f(x,y,z)$ *and its region* R *of definition, a finite closed admissible region, fulfill the following conditions:*

(1) *The function* f *is harmonic* ‖ *in a region containing* R *in its interior.*

---

* The terminology of analysis situs will be used in the sense as defined by J. W. Alexander in his paper "Combinatorial Analysis Situs," *Transactions of the American Mathematical Society*, Vol. 28 (1926), pp. 301-329. But the term "cycle" will be used in place of his term "closed chain" and the symbol "R_i" in place of his symbol "P^i" for the i-th connectivity number.

† A point (x,y,z) is called a *critical point* of a function $f(x,y,z)$ if all the three partial derivatives of first order of f vanish at the point.

‡ See J. J. Gergen, "Mapping of a General Type of Three-Dimensional Region on a Sphere," *American Journal of Mathematics*, Vol. 52 (1930), pp. 197-198. He has proved from different considerations a particular case of our Theorem 1 and a part of our Theorem 4a for a different region.

§ Tsai-Han Kiang, "On the Critical Points of Non-degenerate Newtonian Potentials," Theorem A, *American Journal of Mathematics*, Vol. 54 (1932), pp. 92-109.

¶ A critical point of f is said to be *degenerate* or *non-degenerate* according as the hessian of f vanishes at the point or not. The function f is said to be a *degenerate* or *non-degenerate function* in a region according as it has or has no degenerate critical point in the region.

‖ A function f is said to be *harmonic at a point* if its partial derivatives of second order are continuous and satisfy Laplace's differential equation throughout some neigh-

(2) *The boundary B of R consists of two sets B' and B'' of closed non-singular analytic equipotential surfaces* * *of f, such that the value c' of f on B' is greater than the value c'' of f on B''.*†

Let M_k $(k=1,2)$ be the number of critical points of the k-th type ‡ *of f in R. Let R_i and R'_i $(i=0,1,2)$ be the i-th connectivity numbers of the complexes R and B'' respectively. Then there exist non-negative integers M_k^+ and M_k^- such that*

$$M_k = M_k^+ + M_k^-, \qquad (k=1,2);$$
$$R_0 - R'_0 = -M_1^-, \quad R_1 - R'_1 = M_1^+ - M_2^-, \quad R_2 - R'_2 = M_2^+.$$

We need also the following two lemmas proved by Kellogg.§

LEMMA A. *Suppose $g(x,y,z)$ is the Green's function for a three-dimensional region D with the pole at an interior point of D. Let c' be a non-critical value* ¶ *of g in D. Then the points of D satisfying $g=c'$ constitute one or more closed non-singular analytic surfaces in D, bounding a non-singular analytic region* ‖ *of points of D satisfying $g \geqq c'$.*

---

borhood of that point. It is said to be ***harmonic in a region*** if it is continuous in the region and harmonic at all interior points of the region.

* A ***regular surface element*** and a ***regular surface*** will be used as defined by Kellogg in his book, *Foundations of Potential Theory*, Berlin (1929). A regular surface element will be said to be ***analytic***, if it admits for some orientation of coördinate axes a representation $z=F(x,y)$ where F is analytic. A closed regular surface will be said to be ***non-singular*** (***non-singular analytic***), if every point of the surface is an interior point of a regular (regular analytic) surface element.

† This condition implies that the normal derivative of f never vanishes on B, and as a consequence of harmonicity of f that the constant c' is greater but the constant c'' is less than the value of f at any interior point of R.

‡ Suppose (x^0, y^0, z^0) be a non-degenerate critical point of f. Let f^0_{xx}, f^0_{xy}, etc. be the partial derivatives of second order of f evaluated at the critical point. By a real non-singular transformation linear in x, y, z, the non-singular quadratic form

$$f^0_{xx}(x-x^0)^2 + f^0_{yy}(y-y^0)^2 + f^0_{zz}(z-z^0)^2 + 2f^0_{xy}(x-x^0)(y-y^0)$$
$$+ 2f^0_{yz}(y-y^0)(z-z^0) + 2f^0_{zx}(z-z^0)(x-x^0)$$

can be reduced to one of the forms:

$$\pm X^2 \quad \pm Y^2 \quad \pm Z^2.$$

If the number of negative signs of the reduced form is i, the critical point is said to be of i-th type.

The critical points of 0-th type and third type are critical points at which f has minimum and maximum values respectively. Since f is harmonic in R and since its normal derivative never vanishes on B, $M_0 = M_3 = 0$.

§ Kellogg, *loc. cit.*, pp. 238-239 and p. 276. For Lemma B note also the proposition (c) on the next page.

¶ The value of g at a point of D will be called a ***critical*** value or a ***non-critical value*** of g according as the point is or is not a critical point of g.

‖ A ***non-singular analytic region*** is a finite region in three-dimensional space bounded by one or more closed non-singular analytic surfaces.

LEMMA B. *In a closed region entirely in the interior of D the Green's function has only a finite number of critical values.*

Now let D be an admissible region with the connectivity numbers $R_0 = 1$, R_1, R_2, $R_3 = 0$, and $g(x, y, z)$ the Green's function for D with the pole at an interior point P of D. From the very definition of R_1 and R_2 there exist in D a set of R_1 1-cycles linearly independent with respect to bounding and a set of R_2 2-cycles linearly independent with respect to bounding. We may assume that these R_1 and R_2 cycles are in the interior of D and do not contain the pole P of the Green's function g. Because g is positive in the interior of D, the values of g on these cycles have a positive lower bound, d say. From Lemma B there is a non-critical value c' of g such that $d/2 > c' > d/3$. Let the region of points of D satisfying $g \geqq c'$ be denoted by N. From Lemma A the region N is a non-singular analytic region bounded by one or more closed non-singular analytic equipotential surfaces B' represented by $g = c'$, lies in the interior of D, and contains the cycles in its interior. Moreover, the region N is a 3-complex.* From the definition of connectivity numbers again, the connectivity numbers of N are 1, $R_1 + a_1$, $R_2 + a_2$, 0, where a_1 and a_2 are non-negative integers.

As in the proof of Lemma 4 in the previous paper by the author, *loc. cit.*, the following facts can be established: (a) For a sufficiently large positive constant c'', the points of N satisfying $g = c''$ constitute a single non-singular analytic equipotential surface B'' enclosing the pole P of g. (b) The bounded closed region bounded by B'' is homeomorphic to a closed 3-cell. (c) In this region g has no critical point.

From (b) above, B'' is homeomorphic with a 2-sphere and hence its connectivity numbers are 1, 0, 1, 0. Let R denote the region obtained from N with the interior of B'' removed. From the connectivity numbers of N, the connectivity numbers of R are evidently 1, $R_1 + a_1$, $R_2 + a_2 + 1$, 0, where a_1 and a_2 are non-negative integers.

In case g is degenerate in D, by definition g has at least one critical point in D. Suppose now g is non-degenerate in the interior of the region D and hence in R. Theorem A can be most conveniently applied to the function $-g$ in R. The conclusion of Theorem A states that, when the numbers of critical points of $-g$ of k-th ($k = 1, 2$) types in R are M_k, there exist non-negative integers M_k^+ and M_k^- such that

$$M_k = M_k^+ + M_k^-, \quad -M_1^- = 0, \quad M_1^+ - M_2^- = R_1 + a_1, \quad M_2^+ = R_2 + a_2.$$

---

* S. S. Cairns, "The Celluar Structure and Approximations of Regular Spreads," *Proceedings of the National Academy of Sciences, U. S. A.*, Vol. 16 (1930), pp. 488-491.

Hence for $-g$ we have

$$M_1 = R_1 + a_1 + M_2^-, \qquad M_2 = R_2 + a_2 + M_2^-.$$

Now it is obvious from the definition of the types of critical points, that a critical point of type 1 or 2 of $-g$ is a critical point of type 2 or 1 of $+g$. Hence in R the function g has at least R_2 critical points of type 1 and R_1 critical points of type 2. The results will be summarized in the following theorem.

THEOREM 1. *Suppose the connectivity numbers R_1 and R_2 of an admissible region are not both zero. Then the Green's function for the region with the pole at an interior point has at least one critical point in the interior of the region.*

If, moreover, the Green's function is non-degenerate in the interior of the region, it has at least R_2 critical points of type 1 and R_1 critical points of type 2 in the interior of the region.

3. *Critical points of Green's functions for admissible regions with the same connectivity numbers as a spherical region but not homeomorphic with it.* Let us consider in the space the simple closed curve Q_0 with continuously turning tangent:

$$y^2 = x^4(1 - x^2), \qquad 0 \leqq x, \qquad z = 0;$$

and the point $(r, 0, 0)$ where r is a sufficiently small positive number, less than $\frac{1}{3}$ say. By revolving about the y-axis through two straight angles the point $(r, 0, 0)$ generates a circle C and the curve Q_0 a locus B'_0 of revolution. The origin lies on B'_0 and will be called a *singular point* of B'_0. The finite closed region T'_0 bounded by B'_0 will be called a singular region. In the interior of T'_0 is the circle C.

We can decompose T'_0 into a 3-complex with C as a 1-subcycle. Through any point of B'_0 there is a sphere tangent to B'_0 and containing no point of $T'_0 - B'_0$. From Poincare's criterion for the Dirichlet problem * there exists a Green's function for the region with the pole at any interior point. Hence T'_0 is an admissible region. Obviously the 3-complex T'_0 has the same connectivity numbers as an ordinary spherical region but is not homeomorphic with it, and the circle C does not bound any 2-complex entirely in the interior of T'_0.

Given any positive integer R_1 we can construct in the following manner an admissible region T_0 with the boundary B_0, which has the same connectivity numbers as an ordinary spherical region but is not homeomorphic with it,

* Kellogg, *loc. cit.*, p. 329.

and which has R_1 circles in $T_0 - B_0$ and linearly independent with respect to bounding in $T_0 - B_0$. Let us start with the singular region T'_0. The curve Q_0 has two horizontal tangents at the two points whose x-coördinates are equal to $(2/3)^{1/2}$. Let us subject the whole space to a reflection in the plane $x = (2/3)^{1/2}$. The singular region obtained from T'_0 and its image, each point being counted only once, has a boundary with two singular points, the origin and its image. If we subject the space to another reflection in the plane through one of the two singular points and perpendicular to the x-axis, we shall get a singular region whose boundary has three singular points. After suitable numbers of reflections of these two kinds we shall get a closed singular region T_0 whose boundary B_0 has R_1 singular points. By the same reflections, from C we get R_1 circles $C_1, C_2, \cdots, C_{R_1}$ in $T_0 - B_0$ and linearly independent with respect to bounding in $T_0 - B_0$. The region T_0 has therefore the desired properties.

The region T_0 is an example of the kind of region D in the following theorem.

THEOREM 2. *Suppose D is an admissible region with the same connectivity numbers as an ordinary spherical region but not homeomorphic with it. Suppose there are R_1 1-cycles of D, which are in the interior (in the sense of point sets) of D and are linearly independent with respect to bounding in the interior of D.*

If $R_1 \neq 0$, then the conclusions of Theorem 1 (put $R_2 = 0$) hold for the Green's function for D with the pole at an interior point P of D.

Proof. From our hypothesis there exist R_1 1-cycles in the interior of D. We may assume that these cycles do not contain the pole P. Let d be a positive lower bound of the values of the Green's function g on these cycles. There is a non-critical value c' of g such that $d/2 > c' > d/3$ (Lemma B). The connected non-singular analytic region N of points satisfying $g \geqq c'$ contains the pole P and the cycles $C_1, C_2, \cdots, C_{R_1}$ in its interior (Lemma A). Since N is a sub-region of the interior of D and since the R_1 cycles are linearly independent with respect to bounding in the interior of D, the R_1 cycles are linearly independent with respect to bounding in N. Hence the first connectivity number of the 3-complex N is at least R_1. On applying Theorem 1 to the Green's function $g - c'$ for N with the pole P, we obtain our theorem.

The following corollary is then obvious.

COROLLARY. *The Green's function for the region T_0 with the pole at an interior point has at least one critical point in the interior of T_0.*

If, moreover, the Green's function is non-degenerate in the interior of T_0, it has at least R_1 critical points of type 2 in the interior of T_0.

4. *The region T_0 as the limit of a sequence $\{T_i\}$ of non-singular regions homeomorphic with a spherical region.* Let us return to the curve Q_0 of § 3. Suppose $\{a_i\}$ $(i = 1, 2, 3, \cdots)$ is a monotonically decreasing sequence of positive numbers, whose limit is zero and whose first number a_1 is sufficiently small, less than $1/3$ say. Let us form by means of this sequence of numbers the following sequence of functions of class C':

$$F_i(x) = \begin{cases} (2x + a_i)(x - a_i)^2, & 0 \leqq x \leqq a_i, \\ 0, & a_i \leqq x; \end{cases}$$

and replace the upper branch

$$y = f(x) = x^2(1 - x^2)^{1/2}, \quad 0 \leqq x, \qquad z = 0,$$

of the curve Q_0 by the curve with continuously turning tangent:

$$y = f(x) + F_i(x), \quad 0 \leqq x, \qquad z = 0.$$

The curve Q'_i thus obtained from the whole of Q_0 is a simple open curve with continuously turning tangent, whose two endpoints are at $(0, 0, 0)$ and $(0, a_i^3, 0)$ and which has distinct horizontal tangents at the endpoints. The locus B'_i of revolution generated by revolving Q'_i about the y-axis through two straight angles is thus a non-singular surface. Let T'_i denote the finite closed non-singular region bounded by B'_i.

Let us subject the whole space to the reflections of § 3. Then as we obtained the region T_0 and its boundary B_0 from T'_0 and B'_0 respectively, we shall obtain from the region T'_i and its boundary B'_i a finite region T_i and its boundary B_i. The regions T_i are obviously admissible regions homeomorphic with an ordinary spherical region.

For $i, m = 1, 2, 3, \cdots$, we have

$$\begin{aligned} f(x) &< f(x) + F_{i+m}(x), & 0 \leqq x < a_{i+m}, \\ f(x) + F_{i+m}(x) &< f(x) + F_i(x), & 0 \leqq x < a_i. \end{aligned}$$

Hence T_{i+1} contains T_i as a sub-region and T_i contains T_0 as a sub-region. As i increases the surface B_i shrinks down to B_0. Any point not belonging to T_0 is not a point of T_i either for all values of i or for sufficiently large values of i. It is in this sense that T_i and B_i will be said to *converge from the exterior* to T_0 and B_0 respectively.

Let us observe the following two important properties of the convergent sequence $\{T_i\}$ and the limit region T_0:

(1) Any point of B_0 is a point of B_i either for all values of i or for

sufficiently large values of i. In particular, the singular points of B_0 are points of B_i for all values of i.

(2) Through any point of B_0 there is a sphere which is tangent to B_i and which contains no point of $T_i - B_i$ either for all values of i or for sufficiently large values of i. In particular, through any singular point of B_0 there is a sphere, which is tangent to B_i and contains no point of $T_i - B_i$ for all values of i.

5. *A sequence of harmonic functions* $\{h_i\}$ *in* T_0. The region T_0 is a sub-region of every region T_i $(i = 1, 2, 3, \cdots)$ and an interior point P of T_0 is an interior point of T_i. Since the Dirichlet problem is possible for T_k $(k = 0, 1, 2, \cdots)$, there exists a unique Green's function g_k for T_k with the pole at P. The function g_k is of the form

$$g_k(x, y, z) = 1/r + h_k(x, y, z),$$

where r is the distance from P to the variable point (x, y, z) of T_k, h_k is harmonic in T_k, and $h_k(x, y, z) = -1/r$ for any point (x, y, z) on the boundary B_k of T_k. As a well-known property of harmonic functions, h_k attains its maximum and minimum values only on the boundary B_k. Hence all the functions h_k are bounded in T_0, namely, by the absolute maximum and minimum values of the function $-1/r$ in the finite closed region bounded by B_0 and B_1.

Let us denote any points of T_0 and B_0 by P_0 and b_0 respectively. Since a Green's function is positive in the interior of its region, we have

$$g_0(b_0) \leqq g_{i+1}(b_0) \leqq g_i(b_0), \qquad (i = 1, 2, 3, \cdots),$$

and consequently

$$h_0(b_0) \leqq h_{i+1}(b_0) \leqq h_i(b_0), \qquad (i = 1, 2, 3, \cdots).$$

From the property of harmonic functions stated above, we infer from these inequalities the following:

$$(1) \qquad h_0(P_0) \leqq h_{i+1}(P_0) \leqq h_i(P_0), \qquad (i = 1, 2, 3, \cdots).$$

Now let us consider the sequence $\{h_i\}$ of harmonic functions defined in T_0. From the relations (1), the sequence converges at any point P_0 of T_0. The limit of the sequence is thus a function $H(x, y, z)$ in T_0. By Harnack's second convergence theorem,* the sequence converges uniformly in any closed region in the interior of T_0 and the limit function H is harmonic in the interior of T_0. Moreover, from the property (1) in § 4 the value of H at any point of B_0 is equal to the value of $-1/r$ at that point.

* Kellogg, *loc. cit.*, p. 263.

Identification of H with h_0. To prove that the function H is identically equal to the function h_0, it is only necessary to prove that H takes on the continuous boundary values $-1/r$ on B_0. For, if H takes on these continuous boundary values, H is the solution of the Dirichlet problem for the region T_0 and for this boundary condition, and by the uniqueness theorem of the problem H is identically equal to h_0.

From the property (2) in § 4, through any point b_0 of B_0 there is a sphere which is tangent to B_0 and B_i and which contains no point of $T_0 - B_0$ and $T_i - B_i$ either for all values of i or for sufficiently large values of i. Let $U(A, b_0)$ be the function harmonic in the exterior of the sphere and taking on the continuous boundary values $-1/r$ on the sphere, where A denotes a variable point not in the interior of the sphere and r the distance from P to A. As we obtained (1), so for any point P_0 of T_0 we obtain the following inequalities,

$$h_0(P_0) \leqq h_i(P_0) \leqq U(P_0, b_0),$$

either for all values of i or for sufficiently large values of i. On taking the limit as i increases indefinitely, we find

$$h_0(P_0) \leqq H(P_0) \leqq U(P_0, b_0).$$

Since

$$U(b_0, b_0) = h_0(b_0) = -1/\overline{Pb_0},$$

$\overline{Pb_0}$ being the distance between P and b_0, and since both $h_0(P_0)$ and $U(P_0, b_0)$ are continuous at $P_0 = b_0$, the function $H(P_0)$ in T_0 is continuous at $P_0 = b_0$. Now b_0 is any point of B_0. Hence the function H is continuous in T_0. Hence the two functions H and h_0 are identically the same.

Expressing this result in terms of the Green's functions, we have the following theorem:

THEOREM 3. *Suppose g_k $(k = 0, 1, 2, \cdots)$ is the Green's function for the region T_k (§§ 3-4) with the pole at the same interior point P of T_0. Then the sequence $\{g_i\}$ $(i = 1, 2, \cdots)$ of functions in $T_0 - P$ converges to g_0 in $T_0 - P$, and the convergence is uniform in any closed region in $T_0 - B_0 - P$.*

6. *Critical points of Green's functions for admissible regions homeomorphic with a spherical region.* In § 3 we have defined an admissible region T_0, which has the same connectivity numbers as an ordinary spherical region but is not homeomorphic with it, and which has R_1 circles $C_1, C_2, \cdots, C_{R_1}$ in its interior linearly independent with respect to bounding in its interior. In

§ 4 we have constructed admissible regions T_i $(i = 1, 2, 3, \cdots)$ homeomorphic with an ordinary spherical region, which form a sequence converging to T_0. We shall prove the following theorem.

THEOREM 4a. *There is a positive integer K such that, for $i \geqq K$, the Green's function g_i for the region T_i with the pole at an interior point of T_0 has at least one critical point in the interior of T_0.*

If, moreover, g_i is non-degenerate in the interior of T_0, it has at least R_1 critical points of type 2 in the interior of T_0.

Proof. Let us denote by u_1 a positive lower bound of the values of g_0 on the R_1 circles $C_1, C_2, \cdots, C_{R_1}$ in the interior of T_0. From Lemma B there exist two non-critical values u' and u'' of g_0 such that

$$u_1 > u' > u'' > 0. \tag{2}$$

Let the two closed non-singular analytic regions in T_0 bounded by the surfaces $g_0 = u'$ and $g_0 = u''$ be denoted by E'_0 and E_0'' respectively (Lemma A). The circles are in the interior of E'_0, the region E'_0 is in the interior of E_0'', and the region E_0'' is in the interior of T_0 (Lemma A).

Let us confine our attention to the region E_0''. Let O denote the interior of a sufficiently small sphere about P. The closed region $E_0'' - O$ is a subregion of $T_0 - B_0 - P$. In $E_0'' - O$ the functions g_k $(k = 0, 1, 2, \cdots)$ are harmonic. Let A be any point of $E_0'' - O$. Just as we obtained (1) so we have now

$$g_i(A) - g_0(A) > 0, \qquad (i = 1, 2, 3, \cdots). \tag{3}$$

Since the sequence $\{g_i\}$ converges uniformly to g_0 in $E_0'' - O$ (Theorem 3), for a small positive constant e there exists a positive integer K such that, for $i \geqq K$,

$$g_i(A) - g_0(A) < e. \tag{4}$$

Let us assume that the positive constant e is so small that

$$u' > u'' + e.$$

From Lemma B there exists a non-critical value u_i of g_i $(i \geqq K)$ in T_i such that

$$u' > u_i > u'' + e. \tag{5}$$

Let the non-singular analytic region of points of T_i satisfying $g_i \geqq u_i$ be denoted by E_i. We shall prove the following two statements: (a) E'_0 is in the interior of E_i; and (b) E_i is in the interior of E_0''.

Let A', A_i, A'' denote any points of the surfaces $g_0 = u'$, $g_i = u_i$, $g_0 = u''$ respectively. Since A' is a point of $E_0'' - O$, from (3) we have in particular

$$g_i(A') > g_0(A') = u', \qquad (i \geqq K).$$

But from the first inequality of (5), we have

$$u' > u_i = g_i(A_i).$$

These two relations give

$$g_i(A') > g_i(A_i).$$

From Lemma A applied to g_i, this inequality shows that any point A' of $g_0 = u'$ is in the interior of E_i. The statement (a) is thus proved.

Since A'' is a point of $E_0'' - O$, from (4) we have in particular

$$g_i(A'') - g_0(A'') < e,$$

or

$$g_i(A'') < u'' + e.$$

From the second inequality of (5) we have

$$u'' + e < u_i = g_i(A_i).$$

These two relations give

$$g_i(A'') < g_i(A_i).$$

From Lemma A applied to g_i, this inequality shows that any point A'' of $g_0 = u''$ is not in the interior of E_i. The statement (b) is thus proved.

Now, since the circles $C_1, C_2, \cdots, C_{R_1}$ are in the interior of E'_0, they are in the interior of E_i from the statement (a). Since the circles are linearly independent with respect to bounding in E_0'' and since E_i is a sub-region of E_0'' from the statement (b), the circles are linearly independent with respect to bounding in E_i. Hence the first connectivity number of E_i is at least R_1. Our theorem then follows at once from Theorem 1.

Corollary. *There exists a closed, finite, non-singular region, homeomorphic with an ordinary spherical region, for which the Green's function with the pole at an interior point of the region has at least one critical point in the interior of the region.*

Let us note that Theorem 4a has been deduced from Theorem 1 on the bases of the Corollary to Theorem 2 and of the two properties of the sequence $\{T_i\}$ stated at the end of § 4. By the same method of proof the following general theorem can be easily established.

Theorem 4. *Suppose D is the region in Theorem 2. Suppose $\{D_i\}$ $(i = 1, 2, 3, \cdots)$ is a sequence of admissible regions homeomorphic with an ordinary spherical region, which converges to D from the exterior and has the two properties stated at the end of* § 4.

If $R_1 \neq 0$, then Theorem 4a *holds when T_i and T_0 there are replaced by D_i and D_0 respectively.*

National University of Peking,
Peiping, China.

ON OPERATIONS PERMUTABLE WITH THE LAPLACIAN.

By Hillel Poritsky.†

1. *Introduction.*‡ Gauss' theorem stating that the arithmetic mean $\int u dS / \int dS$ of a function u which is harmonic on and within a spherical surface S, is equal to its value at the center of S, because of its geometrical appeal and elegant simplicity, must be considered among the most attractive theorems in analysis. Koebe, Bôcher, and others have proved converse forms of Gauss' theorem by showing that this property of harmonic functions completely characterized them. Yet comparatively few results in mathematics seem to have had their starting point in that theorem, even when one considers the field of harmonic functions and functions related to them. This paper is the first one of several papers which grew out of an attempt to extend Gauss' mean value theorem.

In trying to organize the results that were first obtained it was found that they could be coördinated and the proofs simplified by the introduction of a certain linear functional operator A, presently to be defined, which is permutable with ∇^2. The proof of this permutability property, given in § 2, Theorem I, is not unlike some of the proofs that have been given for Gauss' theorem, but the more abstract result obtained can be utilized not merely to prove many laws of the spherical mean which generalize Gauss' mean value theorem, but also to derive many apparently disconnected results in the theory of harmonic and related functions, such as Bessel functions.

To describe the operator A, consider in Euclidean 3-space a concentric family of spherical surfaces imbedded in the region of definition of a given function $u(x, y, z)$, where x, y, z are rectangular Cartesian coördinates for the region. Let $\bar{u}(x, y, z)$ be a new function which is constant along each of the concentric spherical surfaces and is equal to the arithmetic mean of u over that spherical surface. The operation $A(u)$ is the operation which replaces u by $\bar{u}$:

$$A(u) = \bar{u} = \int u dS / \int dS.$$

We shall refer to this operation as "averaging" u over concentric spheres,

† National Research Fellow in Mathematics, 1927-1929.

‡ The results of § 2 first appeared in the author's Ph. D. thesis *Topics in Potential Theory*, Cornell 1927, and are published here for the first time.

and to A as the "averaging" operator. The permutability property of A and ∇^2 is expressed by the equation

$$\nabla^2 [A(u)] = A(\nabla^2 u).$$

This permutability is proved in Theorem I of the following section. The proof covers the n-dimensional case for which A is defined in quite a similar manner.

As the range of application of this result soon became quite extensive, it became of interest to look for other linear functional operators which are also permutable with the Laplacian. Such operators generalizing the operator A in various directions were soon found; they are dealt with in §§ 3-9; we proceed to describe them.

In §§ 3, 4 are considered operators of the form

$$L_k(u) = h_k(\omega) \int u(r, \omega') h'_k(\omega') d\omega';$$

here the integration (as in case of the operator A) is carried out over any one of the spherical surfaces of a concentric family imbedded in the domain of definition of u; r is the radius of the spheres, ω a symbolic variable for polar coördinates specifying orientation of rays through the common center, $d\omega$ the element of solid angle subtended at the center by the surface element dS: $d\omega = dS/r^2$; finally, $h_k(\omega)$, $h'_k(\omega)$ are two surface spherical harmonics of degree k, that is, functions of ω such that $h_k r^k$, $h'_k r^k$ are harmonic polynomials of degree k. It will be noticed that for $k = 0$ h_k, h'_k reduce to constants, and L_k becomes proportional to A. The nature of the operators $L_k(u)$ is rendered clear if we consider the case of two dimensions. Introducing polar coördinates, r, θ with pole at the common center and choosing $h_k = e^{ki\theta}$, $h'_k = e^{-ki\theta}/2\pi$ we find that

$$L_k(u) = e^{ki\theta} \int_0^{2\pi} u(r, \theta') \, e^{-ki\theta'} \, d\theta'/2\pi.$$

The operation L_k is thus of the nature of an operation which replaces u by one of the terms in the θ-Fourier expansion of u for each r. Likewise, for any number of dimensions the operators L_k are seen to replace u by the same type of functions as the terms in the (formal) expansion of u over each member of a family of concentric spherical surfaces in terms of a complete set of surface spherical harmonics of various degrees. Two proofs of the permutability of L_k and ∇^2 are given — one involving Euclidean operations only, the other making use of the second differential operator of Beltrami for the spherical surfaces.

The operators L_k find themselves in a sense generalized in the operators

$L_{k,m}$ of § 5. The relation of $L_{k,m}$ to L_k is sufficiently well illustrated by a particular case $(m=1)$ if in three dimensions instead of considering a family of concentric spheres we start with a family of co-axial circles and "average" u over each circle, that is, replace u over each circle by its arithmetic mean over that circle, or, again, replace u by a term of the type occurring in the Fourier expansion of u over each circle in terms of θ, where θ is the central angle along each circle measured from a common half plane through the axis.

In § 6 we consider operators of the form

$$h(x_1, x_2, \cdots, x_m) \int \overset{+\infty}{\underset{-\infty}{\cdots}} \int h'(x'_1, x'_2, \cdots, x'_m)$$
$$\times u(x'_1, x'_2, \cdots, x'_m;\ x_{m+1}, \cdots, x_n)\, dx'_1 \cdots dx'_m.$$

Here the integration is carried out over parallel m-flats immersed in a Euclidean n-space, E_n; $h(x_1, \cdots x_m)$, $h'(x_1, \cdots x_m)$ are solutions of the equation

$$\left(\frac{\partial^2}{\partial x_1^{\,2}} + \cdots + \frac{\partial^2}{\partial x_m^{\,2}}\right)(\ ) - k(\ ) = 0,$$

where k is a constant. These operators are suggested by letting the loci of the common centers of the preceding operators $L_{k,m}$ move off to infinity. Thus for $n=3$, $m=1$ the integrations would extend over parallel lines or planes instead of co-axial circles or concentric spheres.

Returning to the operators

$$L_k(u) = h_k(\omega) \int h'_k(\omega') u(r, \omega')\, d\omega'$$

one might attempt to generalize them by letting $h_k(\omega)$, $h'_k(\omega)$ be *non* integer harmonics, that is, functions of ω such that for a non integer k $r^k h_k$, $r^k h'_k$ are harmonic. Such harmonics, however, prove to be no longer single valued functions of ω, but may be single valued over proper Riemann spaces spread over the unit sphere. These generalizations of L_k are considered in § 7. In § 8 we discuss the analogues of the above results for non-Euclidean spaces. Finally, several further extensions of the results of the preceding sections are considered in § 9.

As regards the proofs it may be observed that they are similar and essentially consist in an application of Green's theorem and in changing the order of differentiations and integration. A considerable part of their complexity is due to the singularity of the coördinate system at the common center of the concentric spheres.

In this paper we confine ourselves entirely to the consideration of the permutability of the operators mentioned with the Laplacian (or its proper

generalization for non-Euclidean space). The application of these results we reserve for future papers. We shall, however, illustrate the manner in which these results are applied by considering a harmonic function u; if O is any linear functional operator permutable with ∇^2, then

$$O(\nabla^2 u) = 0 = \nabla^2 [O(u)]:$$

thus $O(u)$ is also harmonic. Thus one may build new harmonic functions from a given one by applying to it any linear functional operator which is permutable with ∇^2.

2. *Permutability of the Laplacian operator* ∇^2 *with the averaging operator* A. We shall consider functions of n real variables, $x_1, x_2, \cdots, x_n$, where x_i are orthogonal coördinates of a point in n-dimensional Euclidean space E_n. The locus of points of E_n which are a constant distance r away from a fixed point we shall call a "sphere" or "spherical surface" (common terms are "$(n-1)$-sphere," "hypersphere") and shall denote it by S. The $(n-1)$-dimensional element of "area" of S we denote by dS and its projection from the center onto a concentric unit sphere by $d\omega$; unless otherwise stated the integrals $\int f dS$, $\int f d\omega$ will extend over the whole of S. The value of $\int d\omega$, the area of a unit sphere, we denote by K_n. Finally, we write dv for the n-dimensional element of volume of E_n.

As explained in § 1 (for the case $n = 3$), by the "arithmetic mean" or "average" of $u(x_1, x_2, \cdots, x_n)$ over a spherical surface S will be understood the quotient $\int u dS / \int dS$, while the operation which consists in replacing u over each of a concentric family of spherical surfaces (lying in the domain of definition of u) by the average of u over that surface will be denoted by A. The resulting function, $A(u)$, depends on r only. In this manner $A(u)$ has been defined for $r > 0$; we extend its definition to $r = 0$ by defining $A(u)$ for $r = 0$, that is, for the center of the family, as the value of u itself at that point.

We shall now prove

THEOREM I. *The operation A of averaging over concentric spherical surfaces is permutable with the Laplacian operation ∇^2, that is,*

$$\nabla^2 [A(u)] = A [\nabla^2 u]$$

provided u is of class C'' (that is, is continuous and possesses continuous derivatives of first and second order) in the closed region bounded by concentric spheres of radii a, b, $0 < a \leqq r \leqq b$, or by the single sphere of radius b, $r \leqq b$, where r is the distance from the common center. These closed regions we denote by $R_{a,b}$, R_b, respectively.

To prove this we shall first show that the derivatives $\partial^2 A(u)/\partial x_i^2$ exist (this is implied in the statement of the theorem) and are continuous. Let $\theta_1, \theta_2, \cdots, \theta_{n-1}, r$ be a system of polar coördinates, where θ_i are constant along rays through the center; we may choose θ_i, for example, as the angles defined by

$$
(1) \qquad
\begin{aligned}
x_1 &= r\cos\theta_1,\\
x_2 &= r\sin\theta_1\cos\theta_2,\\
&\cdots\cdots\cdots\cdots\\
x_{n-1} &= r\sin\theta_1\sin\theta_2\cdots\sin\theta_{n-2}\cos\theta_{n-1},\\
x_n &= r\sin\theta_1\sin\theta_2\cdots\sin\theta_{n-2}\sin\theta_{n-1},
\end{aligned}
$$

where the origin has been put at the center. Using ω as symbolic for the $n-1$ variables θ_i, we write

$$u = u(r,\omega), \quad dS = r^{n-1}\,d\omega, \quad \int dS = r^{n-1}\int d\omega = r^{n-1}K_n.$$

We may write $A(u)$, $r=0$ not excepted, in the form

$$A(u) = \int u(r,\omega)\,d\omega/K_n,$$

and we notice that both $d\omega$ and the limits of integration are independent of r, while $\partial u/\partial r$, $\partial^2 u/\partial r^2$ exist and are continuous in all the variables r, θ_i. Hence dA/dr, d^2A/dr^2 exist, are continuous in r in the closed interval in question, and may be obtained by differentiating under the integral sign. Moreover, the values of the former derivative for $r=0$ is 0, since by differentiating under the integral sign we see that the directional derivatives in two opposite directions cancel each other. From this the existence and continuity of $\partial A(u)/\partial x_i$, $\partial^2 A(u)/\partial x_i^2$ except at the origin follow at once, and we may write for $r \neq 0$

$$
(2) \qquad \frac{\partial A(u)}{\partial x_i} = \frac{dA(u)}{dr}\,\frac{x_i}{r}, \quad \frac{\partial^2 A(u)}{\partial x_i^2} = \frac{d^2A(u)}{dr^2}\,\frac{x_i^2}{r^2} + \frac{dA(u)}{dr}\,\frac{r^2 - x_i^2}{r^3}.
$$

To prove the existence and continuity of the derivatives of $A(u)$ *at* the origin requires somewhat longer considerations. As regards the *existence* of the derivatives at the origin, let U be any space function depending on r only and of class C'' in r for $0 \leqq r \leqq b$. Along the x_i-axis, since $r = |x_i|$ there, we have $\partial U/\partial x_i = \pm\, dU/dr$ according as $x_i \gtrless 0$, while $\partial^2 U/\partial x_i^2 = d^2U/dr^2$ there for $x_i \neq 0$. Therefore at the origin $\partial U/\partial x_i$ will exist or not according as $dU/dr\,|_{r=0}$ vanishes or not. In the former case, moreover, $\partial^2 U/\partial x_i^2$ is seen to exist along the x_i-axis even for $x_i = 0$ and to be equal to $d^2U/dr^2\,|_{r=0}$ there. Applying this to $A(u)$ we see that $\partial A(u)/\partial x_i$, $\partial^2 A(u)/\partial x_i^2$ exist at the origin and

$$\left.\frac{\partial A(u)}{\partial x_i}\right|_{r=0} = 0, \qquad \left.\frac{\partial^2 A(u)}{\partial x_i^2}\right|_{r=0} = \left.\frac{d^2 A(u)}{dr^2}\right|_{r=0}.$$

Now as regards the *continuity* of the derivatives of $A(u)$ at the origin, if in equations (2) we let r approach zero, since $|x_i/r| \leqq 1$, we get

$$\lim_{r\to 0} \frac{\partial A(u)}{\partial x_i} = 0,$$

$$\lim_{r\to 0} \frac{\partial^2 A(u)}{\partial x_i^2} = \lim_{r\to 0} \frac{d^2 A(u)}{dr^2}\,\frac{x_i^2}{r^2} + \frac{dA(u)}{dr}\,\frac{r^2 - x_i^2}{r^3}.$$

Putting the last term in the form

$$\frac{\dfrac{dA(u)}{dr} - \left(\left.\dfrac{dA(u)}{dr}\right|_{r=0}\right)}{r}\;\frac{r^2 - x_i^2}{r^2}$$

we find

$$\lim_{r\to 0} \frac{\partial^2 A(u)}{\partial x_i^2} = \left.\frac{d^2 A(u)}{dr^2}\right|_{r=0}.$$

Thus $A(u)$ is of class C'' at the origin, too.

Incidentally, it follows from the above considerations that if U is a space function which depends on r only and is of class C'' in r for $0 < a \leqq r \leqq b$, then U is of class C'' in x_i in the region $R_{a,b}$, while if U is of class C'' in r for $0 \leqq r \leqq b$, then U need not be even of class C' in R_b, unless $dU/dr\,|_{r=0} = 0$, in which case U will be of class C'' in R_b. The Laplacian of such a function will be given by

$$\nabla^2 U = \frac{d^2U}{dr^2} + \frac{n-1}{r}\,\frac{dU}{dr} = r^{n-1}\,\frac{d}{dr}\left(r^{n-1}\,\frac{dU}{dr}\right) \quad \text{for} \quad r > 0, \tag{3}$$

$$\nabla^2 U = n\,\frac{d^2U}{dr^2} \quad \text{for} \quad r = 0. \tag{3'}$$

We now proceed with the proof by applying Gauss' Theorem to the spherical shell bounded by two spheres S_1, S_2 of radii $r_1, r_2, 0 < a \leqq r_1 < r_2 \leqq b$:

$$\int \nabla^2 u\, dv = \int \frac{\partial u(r_2,\omega)}{\partial r}\, dS_2 - \int \frac{\partial u(r_1,\omega)}{\partial r}\, dS_1$$
$$= r_2^{\,n-1} \int \frac{\partial u(r_2,\omega)}{\partial r}\, d\omega - r_1^{\,n-1} \int \frac{\partial u(r_1,\omega)}{\partial r}\, d\omega.$$

Dividing by $r_2 - r_1$, letting one of the radii approach the other, replacing the latter by r, and denoting the corresponding spherical surface by S we get

$$(4)\qquad \int \nabla^2 u(r,\omega)\,dS = \frac{d}{dr}\left(r^{n-1}\int \frac{\partial u(r,\omega)}{\partial r}\,d\omega\right).$$

From this follows upon dividing by $\int dS = K_n r^{n-1}$

$$A(\nabla^2 u) = r^{1-n}\,\frac{d}{dr}\left[r^{n-1}\int \frac{\partial u(r,\omega)}{\partial r}\,\frac{d\omega}{K_n}\right].$$

But if we interchange the order of integration and one differentiation and take account of (3), we find that the right hand member above reduces to $\nabla^2 A(u)$. The theorem is thus proved except for $r = 0$. In this case it follows from the continuity of both functions $A[\nabla^2 u]$, $\nabla^2[A(u)]$.

3. *Permutability of the operators L_k with the Laplacian.* We shall now extend the results of § 2 by showing that the property there proved for the averaging operator (namely, its permutability with the Laplacian) is also possessed by the operators L_k defined as follows:

Let $H_k(x_1, x_2, \cdots, x_n)$, $H_k'(x_1, x_2, \cdots, x_n)$ be two homogeneous harmonic polynomials of degree k, or in familiar teminology, two (solid) spherical harmonics of degree k, and let $h_k = H_k/r^k$, $h_k' = H_k'/r^k$ be the corresponding "surface" spherical harmonics. The latter are independent of r, being, in fact, polynomials in $\cos\theta_1$, $\sin\theta_1, \cdots, \sin\theta_{n-1}$. Using the notation of the preceding sections we shall write $h_k = h_k(\omega)$, $h_k' = h_k'(\omega)$. Corresponding to two such spherical harmonics of degree k we define $L_k(u)$:

$$(5)\qquad L_k(u) = h_k(\omega)\int h'_k(\omega')\,u(r,\omega')\,d\omega'.$$

The result of applying L_k to a function u is thus a new function which is a product of the surface spherical harmonic h_k by a function of r; for brevity we shall denote the latter by $I(r)$:

$$I(r) = \int h_k(\omega')\,u(r,\omega')\,d\omega'.$$

At the origin we define $L_k(u)$ for $k \neq 0$ as equal to zero. For $k = 0$ h_k, h_k' are to be replaced by constants and L_0 is seen to be proportional to $A(u)$; this we assume to hold at the origin, too (that is, with the *same* constant of proportionality).

As pointed out in § 1, the functions $L_k(u)$ are of the same type as the functions occurring in the formal expansion of u along each member of the spheres $r = \text{const.}$ in terms of a complete set of surface spherical harmonics independent of r. In this connection the function $I(r)$ appears to be a "Fourier constant" for each spherical surface and the function $L_0(u)$ corresponds to the constant term of the above expansion.

For these L_k we now state

THEOREM II. *The operators L_k are permutable with ∇^2, that is,*

$$\nabla^2 [L_k(u)] = L_k [\nabla^2 u] \tag{6}$$

under the same conditions on u as in Theorem I.

Since Theorem I is contained in Theorem II as a special case and the proof which follows could be rendered independent of the preceding section, the latter could have been omitted. A separate proof of Theorem I seemed, however, desirable in view of its simplicity and its importance for applications, as well as in order to break up the not inconsiderable complexity of the subject matter.

We begin with the consideration of $\partial I/\partial x_i$, $\partial^2 I/\partial x_i^2$. Their existence and continuity for $r > 0$ follows from considerations similar to those of the preceding section. Hence $L_k(u)$ is of class C'' except possibly at the origin.

We proceed to compute $\nabla^2 L_k(u)$:

$$\begin{aligned}\nabla^2 [L_k(u)] &= \nabla^2(h_k I)\\ &= \nabla^2(H_k I r^{-k})\\ &= \nabla^2 H_k(Ir^{-k}) + 2\nabla H_k \cdot \nabla(Ir^{-k}) + H_k \nabla^2(Ir^{-k}),\end{aligned}$$

where the second term is twice the scalar product of the gradients ∇H_k, $\nabla(Ir^{-k})$; this term may be replaced by $2\{(\partial H_k/\partial r)\,[d(Ir^{-k})/dr]\}$ since the gradient of Ir^{-k} points in a radial direction. On equating $\nabla^2 H$ to zero, replacing H_k by $r^k h_k$ and ∇^2 in the last term by (3), and carrying out the differentiations we get

$$\nabla^2 [L_k(u)] = h_k \left[\frac{d^2 I}{dr^2} + \frac{n-1}{r}\,\frac{dI}{dr} + \frac{k(2-n-k)}{r^2}\, I\right]. \tag{7}$$

Next consider

$$L_k(\nabla^2 u) = h_k(\omega)\int h'_k(\omega')\,\nabla^2 u(r,\omega')\,d\omega'.$$

By applying Green's Theorem to the volume between two concentric spherical surfaces which are allowed to approach each other we obtain a result generalizing equation (4):

$$\int (v\nabla^2 u - u\nabla^2 v)\,dS = (d/dr)\,[r^{n-1} \int (v\,\partial u/\partial r - u\,\partial v/dr)\,d\omega]; \tag{8}$$

here v as well as u is of class C''. If in this equation we put H_k' in place of v, we get

$$\int H'_k \nabla^2 u\, dS = (d/dr)\,[r^{n-1} \int (H'_k \partial u/\partial r - u\partial H'_k/\partial r)\,d\omega],$$

and, replacing H'_k by $h'_k r^k$ and dS by $r^{n-1}\,d\omega$,

$$r^{k+n-1}\int h'_k\nabla^2 u\,d\omega = (d/dr)\{r^{n+k-1}[\int h'_k(\partial u/\partial r)\,d\omega - (k/r)\int h'_k u\,d\omega]\};$$

finally, changing the order of differentiation and integration in $\int h'_k(\partial u/\partial r)\,d\omega$, solving for $\int h'_k\nabla^2 u\,d\omega$, and simplifying, we obtain for this integral the bracket of the right-hand member of (7). The proof of Theorem II, except for $r = 0$, is thus complete.

For $r = 0$ we may verify the theorem directly for the cases where u is a polynomial of the second degree in $x_1, \cdots x_n$; it remains to prove it for functions u of class C'' and such that for small r

$$u = o(r^2), \quad \partial u/\partial x_i = o(r), \quad \partial^2 u/\partial x_i^2 = o(1).$$

First we point out that since $h_k = O(1)$,

$$I = \int h'_k(\omega')u(r,\omega')\,d\omega' = o(r^2),$$

$$dI/dr = \int h'_k(\omega')[\partial u(r,\omega')/\partial r]\,d\omega' = o(r), \quad d^2I/dr^2 = o(1);$$

hence

$$\frac{\partial I}{\partial x_i} = \frac{dI}{dr}\,\frac{x_i}{r} = o(r), \quad \frac{\partial^2 I}{\partial x_i^2} = \frac{dI}{dr}\,\frac{r^2 - x_i^2}{r^3} + \frac{d^2I}{dr^2}\,\frac{x_i}{r} = o(1).$$

Again, $\partial h_k(\omega)/\partial x_i = O(r^{-1})$, $\partial^2 h_k(\omega)/\partial x_i^2 = O(r^{-2})$, hence $L_k(u) = h_k I = o(r^2)$ and $L_k(u)$ is continuous at the origin; its partial derivatives with respect to x_i vanish there, for

$$\partial L_k(u)/\partial x_i\,|_{r=0} = \lim_{r\to 0} O(r^2)/x_i = \lim_{r\to 0} O(r) = 0.$$

Near the origin

$$\partial L_k(u)/\partial x_i = (\partial h_k/\partial x_i)I + h_k\partial I/\partial x_i = O(r^{-1})o(r^2) + O(1)o(r) = o(r);$$

$\partial L_k(u)/\partial x_i$ are thus continuous at $r = 0$. Likewise by using the above order relations for the derivatives of I and of h_k, one proves that $\partial^2 L_k(u)/\partial x^2_i$ are continuous at $r = 0$ and vanish there. The last statement also applies, however, to $L_k(\nabla^2 u)$ since

$$L_k(\nabla^2 u) = h_k\int h'_k\nabla^2 u\,d\omega = o(1).$$

We have thus shown that both members of (6) are continuous at the origin and are equal to each other there, too. The proof of Theorem II is thus complete.

The cross derivatives $\partial^2 L_k(u)/\partial x_i\,\partial x_j$ may be shown to be continuous inclu-

sive of the origin in the same manner as was done for $\partial^2 L_k(u)/\partial x_i^2$. Hence $L_k(u)$ is of class C''.

4. *Another proof of Theorem II; the second differential operator of Beltrami.* The proof of Theorem II admits (for $r > 0$) of another form which is illuminating, and reveals the *raison d'être* of the theorem.

Recall expression (3) for ∇^2. Replacing the r- differentiations in it by partial differentiations, we shall denote the result by D_2:

$$(9) \qquad D_2(u) = \frac{\partial^2 u}{\partial r^2} + \frac{n-1}{r}\frac{\partial u}{\partial r} = r^{1-n}\frac{\partial}{\partial r}\left(r^{n-1}\frac{\partial u}{\partial r}\right),$$

and write

$$(10) \qquad \nabla^2 u = D_2 u + \Delta_2 u.$$

The operator Δ_2 is obviously independent of axes and coördinates; the following properties of this operator will be utilized:

1. Δ_2 is a sum of homogeneous differential operators of first and second orders; if polar coördinates r, θ_i are used, Δ_2 involves θ_i—differentiations only.

2. For any two functions $u(\omega)$, $v(\omega)$ which are single valued over a sphere and of class C'' the equation holds

$$(11) \qquad \int [v\Delta_2 u - u\Delta_2 v]\, d\omega = 0.$$

3. A surface spherical harmonic of degree k, h_k, satisfies the equation

$$(12) \qquad \Delta_2(h_k) = -k(k+n-2)h_k r^{-2}.$$

Granting these properties, we shall now prove that ∇^2 and L_k are permutable by showing that L_k is permutable with each of the operators D_2, Δ_2 whose sum is equal to ∇^2. Indeed,

$$D_2[L_k(u)] - L_k(D_2 u) = r^{1-n}(\partial/\partial r\,\{r^{n-1}(\partial/\partial r)[h_k(\omega)\int h'_k(\omega)u(r,\omega')\,d\omega']\}$$
$$-h_k(\omega)\int h'_k(\omega')r^{1-n}(\partial/\partial r\,\{r^{n-1}[\partial u(r,\omega')/\partial r]\}d\omega'$$

and this vanishes since the r-differentiations and the ω-integrations are permutable. Again,

$$\Delta_2[L_k(u)] - L_k(\Delta_2 u) = \Delta_2[h_k(\omega)\int h'_k(\omega')u(r,\omega')\,d\omega']$$
$$-h_k(\omega)\int h'_k(\omega')\Delta_2[u(r,\omega')]\,d\omega'.$$

The right hand member may now be transformed by noticing that in the first term the integral is a function of r only (denoted by $I(r)$ in the preceding section), while Δ_2 involves differentiations along the surface of the sphere only, and utilizing (11); we get

$$\Delta_2 [L_k(u)] - L_k(\Delta_2 u) = [\Delta_2 h_k(\omega)] \int h'_k(\omega') u(r, \omega') d\omega'$$
$$- h_k(\omega) \int u(r, \omega') \Delta_2 [h'_k(\omega')] \, d\omega'.$$

It only remains to make use of (12) to reduce $\Delta_2[L_k(u)] - L_k(\Delta_2 u)$ to zero.

We now turn to the proof of the properties of Δ_2 which have been utilized above. The first one may be proved in a straightforward manner by expressing ∇^2 in terms of r- and θ-differentiations by means of the familiar expression of ∇^2 in curvilinear coördinates. Equation (12) now follows from

$$0 = \nabla^2(H_k) = \nabla^2(h_k r^k)$$

by replacing ∇^2 by the right hand member of (10), noting that D_2 operates only on r^k and Δ_2 only on h_k, and carrying out the r-differentiations involved. Finally, as regards equation (11), it is essentially equivalent to equation (8) of the preceding section. This may be seen by replacing dS in the left hand member of that equation by $r^{n-1} d\omega$, dividing both sides by r^{n-1}, and interchanging the order of the integration and the subsequent operations on the right; we get

$$\int (v\nabla^2 u - u\nabla^2 v)\, d\omega = \int r^{1-n} (\partial/\partial r) [r^{n-1} (v\partial u/\partial r - u\partial v/\partial r)] \, d\omega$$
$$= \int [(n-1)/r] (v\partial u/\partial r - u\partial v/\partial r)\, d\omega + \int \left(\frac{\partial v}{\partial r} \frac{\partial u}{\partial r} - \frac{\partial u}{\partial r} \frac{\partial v}{\partial r} \right) d\omega$$
$$+ \int (v\partial^2 u/\partial r^2 - u\partial^2 v/\partial r^2)\, d\omega$$
$$= \int (vD_2 u - uD_2 v)\, d\omega.$$

Transposing and replacing $\nabla^2 - D_2$ by Δ_2 we obtain the formula in question.

Another method of deducing the properties of the operator Δ_2 is also of interest. This operator constitutes for the sphere what is known as the "second differential invariant operator", due to Beltrami. For any Riemannian space with a metric

$$ds^2 = \sum_{i,j} g_{ij} d\xi_i d\xi_j \qquad (i, j = i, \cdots, n),$$

the latter operator might be defined by

(13) $\Delta_2(u) = g^{-1/2} \sum_i (\partial/\partial\xi_i)[g^{1/2} \sum_j g^{ij}(\partial u/\partial\xi_j)] \qquad (i, j = 1, \cdots, n)$;

here g is the absolute value of the determinant $|g_{ij}|$ ($g\, d\xi_i \cdots d\xi_n$ is thus the element of volume $d\tau$ of R_n), while g^{ij} is the matrix which is "reciprocal" to the matrix g_{ij}. In particular, if ξ_i are orthogonal coördinates, so that

$$ds^2 = \sum h_i^2\, d\xi_i^2 \qquad (i = 1, \cdots, n);$$

(13) becomes

$$(13') \qquad \Delta_2 = \frac{1}{h_i \cdots h_n} \sum_i \frac{\partial}{\partial\xi_i}\left(\frac{h_1 \cdots h_{i-1}\, h_{i+1} \cdots h_n}{h_i} \frac{\partial u}{\partial\xi_i}\right).$$

This operator may be proved to possess a "Green" theorem similar to that possessed by the Laplacian ∇^2 (to which it reduces if R_n is reduced to E_n):

$$(14) \qquad \int (v\Delta_2 u - u\Delta_2 v)\, d\tau = \int [v(\partial u/\partial n) - u(\partial v/\partial n)]\, d\Sigma,$$

where the left hand integral extends over a region τ of R_n and the right hand integral over its boundary Σ.

Now equation (11) constitutes a special instance of this general integration theorem: if for R_n we choose a sphere S in E_n and apply (14) to the complete sphere, the right hand member reduces to zero since there is no boundary, and we get

$$\int (v\Delta_2 u - u\Delta_2 v)\, dS = 0;$$

from this (11) follows by dividing by r^{n-1} provided that the identity of the present operator Δ_2 with the operator Δ_2 as defined by (10) is granted.

The identity of these two operators may be established without resorting to detailed computations. If in (14) we put $v = 1$, it reduces to

$$(14') \qquad \int \Delta_2 u\, d\tau = \int (\partial u/\partial n)\, d\Sigma.\dagger$$

† From this one deduces the well known definition of Δ_2 as the limit of the ratio $\int (\partial u/\partial n)\, d\tau/\tau$ as τ shrinks to a point; the invariant nature of Δ_2 as regards coördinates is thus manifest.

It will be recalled that the first invariant differential parameter Δ_1, due to Lamé, is the maximum of the various directional derivatives of u at a point. A further intrinsic definition of Δ_2 is through the variation of the integral $\int \Delta_1 u\, dt$, thus:

$$\delta \int \Delta_1(u)\, d\tau = \int \Delta_2 u\, \delta u\, d\tau$$

for variations δu of u, which vanish on the boundary of τ. See, for instance, Courant Hilbert, *Methoden der Mathematischen Physik*, Vol. 1, p. 194; also W. Blaschke, *Vorlesungen über Differentialgeometrie*, Vol. 1, §§ 66-68.

Next apply Gauss' theorem $\int \nabla^2 u\, dv = \int (\partial u/\partial n)\, dS$ to the volume bounded by two concentric spherical surfaces $r = r_1$, $r = r_2$, and a solid cone obtained by letting ω range over a proper region ω_R of the unit sphere. If we divide by $(r_2 - r_1)$ and let r_1, r_2 approach r, the left-hand member approaches $\int_{S_R} \nabla^2 u dS$, where S_R is the portion of S for which ω lies in ω_R; the contribution to the surface integral from the two concentric surfaces, upon dividing by $r_2 - r_1$, is seen to approach

$$\frac{d}{dr}\left(r^{n-1}\int_{\omega_R}\frac{\partial u(r,\omega)}{\partial r}\, d\omega\right) = \int_{\omega_R}\frac{\partial}{\partial r}\left[r^{n-1}\frac{\partial u(r,\omega)}{\partial r}\right] d\omega$$

$$= \int_{S_R} r^{1-n}\frac{\partial}{\partial r}\left[r^{n-1}\frac{\partial u(r,\omega)}{\partial r}\right] dS = \int_{S_R}\left(\frac{\partial^2 u}{\partial r^2} + \frac{n-1}{r}\frac{\partial u}{\partial r}\right) dS;$$

finally, the contribution to the surface integral from the boundary of the solid angle gives rise to a similar $[(n-2)$-dimensional$]$ integral over the boundary of S_R, involving derivatives of u in directions normal to this boundary and tangent to S; by means of (14′) this integral may be changed to $\int_{S_R} \Delta_2 u\, dS$. Hence we get

$$\int_{S_R} \nabla^2 u\, dS = \int_{S_R} \{\partial^2 u/\partial r^2 + [(n-1)/r](\partial u/\partial r)\} dS + \int_{S_R} \Delta_2 u\, dS,$$

and on equating the integrands obtain equation (10). We have thus established the identity of the two ways of introducing Δ_2.

5. *Extension of the property of permutability with the Laplacian to the operators $L_{k,m}$.* We now consider a Euclidean space E_{n+m} of $n+m$ dimensions with the rectangular coördinates $x_1, \cdots, x_n;\ x_{n+1}, \cdots, x_{n+m}$, and a function $u(x_1, \cdots, x_n;\ x_{n+1}, \cdots, x_{n+m})$ of class C''. If we equate $x_{n+1}, \cdots, x_{n+m}$ to the constants $C_{n+1}, \cdots, C_{n+m}$, we find ourselves in a Euclidean space E_n of n dimensions. To the resulting function $u(x_1, \cdots, x_n;\ C_{n+1}, \cdots, C_{n+m})$ we may apply the operators L_k of the last section, obtaining thereby a new function over E_n. If we now vary the constants $C_{n+1}, \cdots, C_{n+m}$ while we keep fixed the harmonics H_k, H'_k in terms of which L_k is defined † and transform u into $L_k(u)$ over each of the E_n thus obtained, there results a new function over E_{n+m}; this function we denote by $L_{k,n}(u)$. We shall now show that these operators $L_{k,n}$ are permutable with

† These harmonics involve $x_1, \cdots, x_n$ only.

$$\nabla^2 = \partial^2/\partial x_1^2 + \partial^2/\partial x_2^2 + \cdots + \partial^2/\partial x^2_{n+m}.$$

Indeed, by Theorem II $L_{k,n}$ is permutable with $\partial^2/\partial x_1^2 + \cdots + \partial^2/\partial x_n^2$;

it remains to show that it is also permutable with $\partial^2/\partial x^2_{n+1} + \cdots + \partial^2/\partial x^2_{n+m}$. Now for $r > 0$ L_k is given by

$$(5) \qquad L_k(u) = h_k(\omega) \int h'_k(\omega')u(r, \omega'; x_{n+1}, \cdots, x_{n+m})\, d\omega'$$

where $r^2 = x_1^2 + \cdots + x_n^2$ and ω stands for $\theta_1, \cdots, \theta_{n-1}$ defined—say—by equations (1). Thus both $h_k(\omega)$, $h'_k(\omega')$ as well as the limits of integration are independent of $x_{n+1}, \cdots, x_{n+m}$; because of this and since $\partial u/\partial x_{n+1}$ is continuous in all the variables $r, \theta_1, \cdots, \theta_{n+1}$; $x_{n+1}, \cdots, x_{n+m}$, it follows that $\partial L_{k,n}/\partial x_{n+1}$ exists and may be obtained by differentiating under the integral sign; there, of course, the differentiation is applied to u only; we have thus shown that for $r > 0$ $L_{k,n}$ is permutable with $\partial/\partial x_{n+1}$; likewise it is proved to be permutable with $\partial/\partial x_{n+2}, \cdots, \partial^2/\partial x^2_{n+1}, \partial^2/\partial x^2_{n+2}, \cdots$, and hence with $\partial^2/\partial x^2_{n+1} + \cdots + \partial^2/\partial x^2_{n+m}$.

Now consider the locus $r = 0$. Denote the operator $\partial^2/\partial x^2_{n+1} + \cdots + \partial^2/\partial x^2_{n+m}$ by $\bar{\nabla}^2$. The function $\bar{\nabla}^2 u$ is continuous inclusive of $r = 0$. Hence we infer from Theorem II that $L_k(\bar{\nabla}^2 u)$ is continuous in $x_1, \cdots, x_n$ at each point of $r = 0$. Along the latter locus, however, $L_k(\bar{\nabla}^2 u)$ reduces by definition either to zero ($k > 0$) or to a constant multiple of $\bar{\nabla}^2 u$ ($k = 0$), and is thus continuous *within* $r = 0$. $L_k(\bar{\nabla}^2 u)$ is therefore continuous at $r = 0$. Likewise $L_k(u)$ reduces for $r = 0$ either to zero or to the same constant multiple of u. Therefore $\bar{\nabla}^2 L_k(u)$, already proved existent and equal to $L_k(\bar{\nabla}^2 u)$ for $r > 0$, is seen to exist at $r = 0$ and is equal to $L_k(\bar{\nabla}^2 u)$ there, too.

In this way is proved the permutability of $L_{k,n}$ with ∇^2 for regions given by

$$a(x_{n+1}, \cdots, x_{n+m}) \leqq r \leqq b(x_{n+1}, \cdots, x_{n+m}),$$

where a, b are continuous functions of $x_{n+1}, \cdots, x_{n+m}$. For more general regions of E_{n+m}, obtained by letting $r, x_{n+1}, \cdots, x_{n+m}$ range over an $(m+1)$-dimensional region, the proof may be carried out by approximating to the boundary by means of a finite number of regions whose boundaries are of the above type. For brevity we shall refer to such regions (and to regions obtained from them by a rigid movement of space) as "*regions over which $L_{k,n}$ is applicable.*"

If we now revert to employing n for the *total* number of dimensions of the Euclidean space under consideration and replace the above n by m ($\leqq n$), we may summarize the above result in a theorem which forms an extension of Theorem II:

THEOREM III. *The operators $L_{k,m}$ are permutable with the Laplacian, that is,*

$$\nabla^2[L_{k,m}(u)] = L_{k,m}(\nabla^2 u)$$

provided u is of class C'' in a region over which $L_{k,m}$ is applicable.

6. *Permutability of the Laplacian with the operators $L^*_{k,m}$.* We shall now consider the operators $L^*_{k,m}$ defined by

$$(15) \quad L^*_{k,m}(u) = h_k(x_1, x_2, \cdots, x_m) \int \overset{+\infty}{\underset{-\infty}{\cdots}} \int h'_k(x'_1, x'_2, \cdots, x'_m)$$
$$\times u(x'_1, x'_2, \cdots, x'_m; x_{m+1}, \cdots, x_n)\, dx'_1, \cdots, dx'_m,$$

where $h_k(x_1, \cdots, x_m)$, $h_k'(x_1, \cdots, x_m)$ are solutions of

$$(\partial^2/\partial x_1^2 + \cdots + \partial^2/\partial x_m^2) h = kh$$

for a constant k for $-\infty < x_1, \cdots, x_m < +\infty$.

It is of interest to point out (though unnecessary for the proof which follows) that these operators may be considered as the limits approached by the operators L_k, $L_{k,m}$ of the preceding sections when the centers of the spheres or subspheres recede to infinity and the degree of the surface harmonics h_k, h'_k is properly increased. Thus, if for definiteness, we consider in E_3 a concentric family of spheres with an associated operator

$$(5) \qquad L_k(u) = h_k(\omega) \int h'_k(\omega') u(r, \omega')\, d\omega'$$

and let the center recede to infinity, say, in the direction of the x_3-axis, the spheres $r =$ constant flatten out into the planes $x_3 =$ constant. Now it was shown in § 4 that along a sphere of radius r surface harmonics $h_k(\omega)$ of degree k satisfy the equation,

$$(12) \qquad \Delta_2 h_k = -k(k+n-2) r^{-2} h_k$$

(with $n = 3$). Therefore, if at the same time that the center recedes to infinity we let k become infinite so that $-k(k+n-2)r^{-2}$ approaches a finite limit k', we are led to consider the equation

$$(\partial^2/\partial x_1^2 + \partial^2/\partial x_2^2) h + k'h = 0$$

as the limit of (12), and the operator L^*_k as the limit of L_k.

A familiar instance of the above limiting process is the passage from the Legendre polynomials to Bessel's function:

$$\lim_{n\to\infty} P_n(\cos\theta/n) = J_0(\theta).\dagger$$

† See, for instance, G. N. Watson, *Bessel Functions*, p. 155.

Returning to the general operator $L^*_{k,m}$ defined by (15), we now state:

THEOREM IV. *The operators $L^*_{k,m}$ are permutable with ∇^2 provided that*

1. *u is of class C'' in any region R which is the product complex of any finite region R_m in the $(x_1, x_2, \cdots, x_m)$—space by a finite region R_{n-m} in the $(x_{m+1}, \cdots, x_n)$—space;*

2. *The integrals*

$$I_1 = \int \overset{+\infty}{\underset{-\infty}{\cdots}} \int h'_k(x_1, \cdots, x_m) u(x_1, \cdots, x_n)\, dx_1, \cdots, dx_m,$$

$$I_2 = \int \overset{+\infty}{\underset{-\infty}{\cdots}} \int h'_k(x_1, \cdots, x_m)(\partial^2/\partial x^2_{m+1} + \cdots + \partial^2/\partial x_n^2) u(x_1, \cdots, x_n)\, dx_1 \cdots dx_n$$

$$I_3 = \int \overset{+\infty}{\underset{\infty-}{\cdots}} \int h'_k(x_1, \cdots, x_m)(\partial^2/\partial x_1^2 + \cdots + \partial^2/\partial x_m^2) u(x_1, \cdots, x_n)\, dx_1 \cdots dx_m$$

converge.

3. *The order of the differentiations and integrations occurring in the integral I_2 may be inverted, that is,*

$$(16)\quad (\partial^2/\partial x^2_{m+1} + \cdots \partial^2/\partial x_n^2) \int \overset{-\infty}{\underset{+\infty}{\cdots}} \int h'_k(x_1, \cdots, x_m) u(x_1, \cdots, x_n)\, dx_1 \cdots dx_m$$

$$= \int \overset{+\infty}{\underset{-\infty}{\cdots}} \int h'_k(x_1, \cdots, x_m)(\partial^2/\partial x^2_{m+1} + \cdots + \partial^2/\partial x_n^2) u(x_1, \cdots x_n)\, dx_1 \cdots dx_m.\dagger$$

4. *The integral*

$$\int_{S_{m-1}} (u\, \partial h'_k/\partial n - h'_k\, \partial u/\partial n)\, dS_{m-1}$$

extended over the boundary S_{m-1}, of the region R_m, where $\partial/\partial n$ denotes the derivative in the direction of the outer normal, approaches zero as R_m expands to infinity, that is, as it expands so as to enclose any point of the $x_1, \cdots, x_m$-space.

To prove this break up the operator ∇^2 into two parts:

$$\nabla^2 = (\partial^2/\partial x_1^2 + \cdots + \partial^2/\partial x_m^2) + (\partial^2/\partial x^2_{m+1} + \partial^2/\partial x_n^2) = {}'\nabla^2 + {}''\nabla^2.$$

We shall show that each part is permutable with $L^*_{k,m}$.

First consider the operator $''\nabla^2$. Conditions for its permutability with

† A sufficient condition for this is that the right hand member of (16) converge uniformly with respect to $x_{m+1}, \cdots, x_n$. See de la Vallée Poussin, *Cours d'Analyse*, Vol. 2, § 23.

$L^*_{k,m}$ are essentially contained in condition 3 above. Thus, equation (16) implies that $''\nabla^2$ may be applied to the integral

$$I_2 = \int_{-\infty}^{+\infty} \cdots \int h'_k(x'_1, \cdots, x'_m) u(x'_1, \cdots, x'_m;\ x_{m+1}, \cdots, x_n)\, dx'_1 \cdots dx'_m;$$

hence it may also be applied to $h_k(x_1, \cdots, x_m)\, I_2$ yielding

$$''\nabla^2 (h_k I_2) = h_k\, ''\nabla^2 I_2.$$

Multiplying both members of (16) by h_k and transforming the resulting left hand member by means of the last equation, we obtain

$$''\nabla^2 L^*_{k,m}(u) = L^*_{k,m}(''\nabla^2 u).$$

Next consider the operator $'\nabla^2$.

$$\begin{aligned} L^*_{k,m}('\nabla^2 u) &= h_k(x_1, \cdots, x_m) \int_{+\infty}^{-\infty} \cdots \int h'_k(x'_1, \cdots, x'_m) \\ &\quad \times (\partial^2/\partial x'^2_1 + \cdots + \partial^2/\partial x'^2_m) u(x'_1, \cdots, x'_m;\ x_{m+1}, \cdots, x_n)\, dx'_1 \cdots dx'_m \\ &= h_k \lim_{R_m \to \infty} \int_{R_m} h'_k\, '\nabla^2 u\, dv_m, \end{aligned}$$

where the notation is obvious; the integral involved is precisely I_3; it may be transformed by means of Green's theorem thus:

$$\int_{R_m} h'_k\, '\nabla^2 u\, dv_m = \int_{R_m} ('\nabla^2 h'_k) u\, dv_m + \int_{S_{m-1}} (h'_k\, \partial u/\partial n - u\, \partial h'_k/\partial n)\, dS_{m-1}.$$

Now on account of condition 4, the last integral approaches zero as R_m expands to infinity. Hence,

$$\begin{aligned} L^*_{k,m}('\nabla^2 u) &= h_k \lim_{R_m \to \infty} \int_{R_m} ('\nabla^2 h'_k) u\, dv_m \\ &= kh_k \lim_{R_m \to \infty} \int_{R_m} h'_k\, u\, dv_m \\ &= kh_k I_1. \end{aligned}$$

On the other hand,

$$'\nabla^2 L^*_{k,m}(u) = '\nabla^2 (h'_k I_1) = ('\nabla^2 h'_k) I = kh'_k I_1$$

since I_1 is independent of $x_1, \cdots, x_m$. Hence $L^*_{k,m}$ is permutable with the operator $'\nabla^2$ as well. The proof of Theorem IV is thus complete.

7. *Generalization to multiple-valued harmonics.* Returning, say, to the operators L_k of §§ 3, 4, one may note on examining the proof of Theorem II

that while use is made of the fact that $h_k(\omega)r^k$, $h'_k(\omega)r^k$ are harmonic, no explicit use is made of the facts that k is an integer, and that these functions are polynomials in $x_1, \cdots, x_n$. The proof, it appears, would go thru just as well if for a *non*-integer constant k the functions $h_k(\omega)r^k$; $h'_k(\omega)r^k$ were harmonic, or, what amounts to the same thing, if for such a constant k the functions $h_k(\omega)$, $h_k'(\omega)$ satisfied the equation (12)

$$\Delta_2 h_k = -k(k+n-2)h_k .$$

along the unit sphere. However, such an extension of Theorem II, while possible, is really vacuous due to the circumstance that the only values of the parameter κ for which there exists a non-zero solution of

$$\Delta_2 h = \kappa h$$

along a unit sphere (in E_n) are precisely the values $\kappa = -k(k+n-2)$, where κ is an integer, and that for any such characteristic parameter value the function hr^k is a harmonic polynomial of degree k. For this reason the direct extension of Theorem II to non-integer k is impossible.

To get an idea of how the above results may be extended to more general surface harmonics, that is, to general solutions of (12) for which $r^k h_k$ are not polynomials, consider the case $n=2$. In this case equation (12) reduces along the unit circle to $d^2h/d\theta^2 = -k^2h$; the solutions of the latter equations are linear combinations of $e^{ki\theta}$, $e^{-ki\theta}$ and are thus single-valued over the unit circle if and only if k is an integer. For other values of k these solutions may be considered to be single-valued over a proper closed or open curve Ω which covers the unit circle a finite or infinite number of times. The operator

$$h_k(\theta)\int h'_k(\theta')u(r,\theta')\,d\theta'$$

may now be applied to any function u which is similarly multiple valued for each r; that is, single-valued over the Riemann surface which is the product of Ω by an r-interval. The proof of Theorem II may be easily adapted to the case of rational k.

It will be recalled that in the proof of § 4, ∇^2 is broken up into D_2 and Δ_2, and each of these is shown to be permutable with L_k. The permutability of D_2 follows from a change of the order of its r-differentiations and ω-integrations; that of Δ_2 from equation (11), itself a special case of the Green-Beltrami Theorem (14). For $n=2$, we have

$$D_2 = \partial^2u/\partial r^2 + (1/r)(\partial u/\partial r), \qquad \Delta_2 = \partial^2 u/r^2\partial\theta^2,$$

and (11), (14) reduce to

$$\int_0^{2\pi}(v\,\partial^2u/r^2\partial\theta^2 - u\,\partial^2v/r^2\partial\theta^2)\,d\theta = 0,$$

$$\int_{\theta_1}^{\theta_2}(v\,\partial^2u/r^2\partial\theta^2 - u\,\partial^2v/r^2\partial\theta^2)\,d\theta = v\,\partial u/r\partial\theta - u\,\partial v/r\partial\theta\Big]_{\theta_1}^{\theta_2}$$

respectively.

It will be seen that for rational k, $k = h/l$, where h and l are relatively prime integers, the proof applies provided that the limits 0 and 2π are placed by 0 and $2\pi l$. However, for irrational k, the operator in question becomes

$$h_k(\theta)\int_{-\infty}^{\infty}h'_k(\theta')u(r,\theta')\,d\theta',$$

and proper additional conditions have to be introduced due to the infinite limits; this may be done in a fashion quite analogous to that of the preceding section in connection with $L_{k,m}$ for which the range of integration was infinite.

In quite a similar manner one could treat for any n the operator

$$\Delta_k(u) = h_k(\Omega)\int h'_k(\Omega')u(r,\Omega')\,d\Omega',$$

where h_k, h_k' are *any* two solutions of the equation (12)

$$\Delta_2 h = kh$$

over the unit sphere, and Ω, Ω' are the Riemann surfaces spread over the unit sphere over which h_k, h_k' are respectively single-valued; the function u is single-valued over the product complex of Ω' by an r-interval, $0 < a \leqq r \leqq b$. For $n = 2$, Ω, Ω' necessarily possess the same structure; such need not be the case, however, for $n > 2$. The Riemann surfaces Ω, Ω' are not to be confused with the spaces of Riemann differential geometry but are to be thought of as consisting of a finite or infinite number of possibly incomplete copies of the unit sphere, properly cut and cross-connected. Points of Ω' at which h'_k is non-analytic along the unite sphere will be considered as belonging to the boundary of Ω'; this boundary will include points in whose neighborhood h'_k fails to be single-valued along the unit sphere. Points not on the boundary of Ω' will be called "inner" points.

The conditions for the permutability of Δ_k and ∇^2, formulated below in Theorem V, are quite analogous to those of Theorem IV of the preceding section. The analogue of the finite region R_m is taken by a part Ω'_f of Ω', which consists of a finite number of possibly incomplete copies of the unit sphere, and contains no boundary point of Ω' either within it or on its boundary; we shall call Ω'_f a "finite part" of Ω'.

To avoid unnecessary repetitions Theorem V is stated for the operators $\Delta_{k,m}$ defined by

$$\Delta_{k,m}(u) = h_k(\Omega) \int h'_k(\Omega') u(r, \Omega'; x_{m+1}, \cdots, x_n) \, d\Omega';$$

and which reduce to the operators just described, but for an m-dimensional case, by putting $x_{m+1}, \cdots, x_n$ equal to constants. Ω, Ω' are Riemann surfaces spread over the "sub-spheres" $x_1^2 + x_2^2 + \cdots + x_m^2 = r^2$. These operators extend the operators $L_{k,m}$ of § 5 in the same way that Δ_k extend the operators L_k, and reduce to Δ_k for $m = n$.

THEOREM V. *The operators $\Delta_{k,m}$ are permutable with ∇^2 except where $h_k(\Omega)$ is not of class C'', provided that*

1. *u is of class C'' in the product complex of Ω' by an r-interval $0 < a \leqq r \leqq b$ by a region in the $(x_{m+1}, \cdots, x_n)$-space.*

2. *The integrals*

$$I_1 = \int_{\Omega'} h'_k(\Omega') u(r, \Omega'; x_{m+1}, \cdots, x_n) \, d\Omega',$$

$$I_2 = \int_{\Omega'} h_k'(\Omega') \{\partial^2/\partial r^2 + [(m-1)/r](\partial^2/\partial r) + \partial^2/\partial x^2_{m+1} + \cdots + \partial^2/\partial x_n^2\} u d\Omega'$$

$$I_3 = \int_{\Omega'} h'_k(\Omega') \Delta_2 u \, d\Omega',$$

if improper, converge.

3. *The order of integrations and differentiations in I_2 may be inverted.*

4. *The integral*

$$\int (u \, \partial h'_k/\partial n - h'_k \, \partial u/\partial n) \, d\Sigma$$

extended over the boundary Σ of a finite part Ω'_f of Ω' approaches zero as Ω'_f expands so as to enclose any inner point of Ω'.

The proof of Theorem V parallels closely the proof of Theorem IV and will be omitted.

It is of interest to point out that Theorem V remains valid if the Ω' integrations are carried out not over all of Ω', but only over part of it, provided that along the additional boundary the integral

$$\int (u \, \partial h'_k/\partial n - h'_k \, \partial u/\partial n) \, d\Sigma$$

vanishes.

Theorem IV, too, admits of similar extensions to multiple-valued functions h_k, h_k' or to integrations extending over part of the flat or the Riemann surface over which h_k' is single-valued.

8. *Non-Euclidean Spaces.* We shall denote the non-Euclidean space by N_n and consider the analogues of the operators L_k, $L_{k,m}$, $L^*_{k,m}$. Naturally, the Laplacian ∇^2 of E_n is to be replaced by the second differential invariant operator of Beltrami Δ^2 for N_n.

We recall first that the geometry of directions thru a point is the same for N_n as for E_n; therefore we may use θ_i, ω to specify directions thru a point, as in the Euclidean case, without leading to any confusion. The element of length in spherical coördinates r, ω is

$$ds^2 = dr^2 + (\sin cr/c)^2 (ds_\omega)^2,$$

where c is the reciprocal of the space constant and ds_ω is the element of length along a unit sphere in E_n; c^2 is positive for Riemann spaces, negative for Lobachevsky spaces, while the Euclidean case is obtained by letting c^2 approach zero. In terms of these spherical coördinates the operators L_k are defined as in the Euclidean case by means of

$$L_k(u) = h_k(\omega) \int h'_k(\omega') u(r, \omega') d\omega';$$

here h_k, h_k' are the *same* surface spherical harmonics as in the Euclidean case. Likewise the operators $L_{k,m}$ are obtained by applying the m-dimensional case of L_k to m-flats which are totally perpendicular to a fixed $(n-m)$-flat, N_{n-m}:

$$L_{k,m}(u) = h_k(\omega) \int h'_k(\omega') u(r, \omega'; \xi_1, \cdots, \xi_{n-m}) d\omega;$$

here $\xi_1, \cdots, \xi_{n-m}$ are coördinates along N_{n-m}, specifying the center of the subspheres of integration, r is the radius of the subspheres, and ω specifies the direction of the radii; directions thru two centers O_1, O_2, corresponding to the same ω are directions perpendicular to the line O_1O_2 and lying in a half plane bounded by it.

To prove the permutability of $L_{k,m}$ with Δ_2 (the operators L_k result from $L_{k,m}$ if $m = n$), we observe that in terms of the coördinates $r, \omega; \xi_1, \cdots, \xi_{n-m}$ the element of length for N_n is given by

$$ds^2 = dr^2 + \cos^2 cr (ds_\xi)^2 + (\sin^2 cr/c^2)(ds_\omega)^2.$$

where $(ds_\xi)^2$ is the quadratic differential in $d\xi_1, \cdots, d\xi_{n-m}$ which is equal to the square of element of length along N_{n-m}, and $(ds_\omega)^2$, similarly, is equal to

the square of element of length along a Euclidean unit sphere in E_m. Choosing for ξ_i as well as for ω orthogonal coördinates, it is readily verified that

$$\Delta_2 u = \frac{c^2}{\cos^{n-m} cr \sin^{m-1} cr} \frac{\partial}{\partial r}\left[\cos^{n-m} cr \sin^{m-1} cr \frac{\partial u}{\partial r}\right]$$
$$+ \frac{1}{\cos^2 cr}\ {}_{\xi}\Delta_2 u + \frac{(cr)^2}{\sin^2 cr}\ {}_{\omega}\Delta_2 u,$$

where ${}_{\xi}\Delta_2$, ${}_{\omega}\Delta_2$ are the same differential operators as the Beltrami operators for N_{n-m}, and for a Euclidean sphere in E_m, respectively.

We may now show that each of the three parts of Δ_2 is permutable with $L_{k,m}$. Indeed, the permutability of the first two parts follows by changing the order of the ω-integrations and the r- or ξ_i-differentiations. The permutability of the third part follows by applying adaptations of equations (11) and (12) in the form,

$$\int v({}_{\omega}\Delta_2 u) - u({}_{\omega}\Delta_2 v)\, d\omega = 0,$$
$${}_{\omega}\Delta_2(h_k, h_k') = -k(k+m-2)(h_k, h'_k).$$

The proof may be extended to $r = 0$ by means of considerations similar to those used in §§ 3, 5 for the Euclidean cases.†

It will be recalled that in the Euclidean case the operators $L_{k,m}$ are suggested by letting the locus of centers recede to infinity. While no such limiting process may be carried out in the Riemann geometry, for the Lobachevsky geometry the locus of centers N_{n-m} could be placed either at infinity or in the transfinite region. The operators $L^*_{k,m}$ which result involve integrations not along spheres, but along horospheres and equidistant surfaces, respectively. For the transfinite case analogue of $L^*_{k,m}$; it is of interest to note, that the same coördinate system is involved as for $L_{k,n-m}$. The operator in question is

$$L^{**}_{k,m}(u) = h_k(\xi_1, \cdots, \xi_m) \int h'_k(\xi'_1, \cdots, \xi'_m) u(r, \omega; \xi'_1, \cdots, \xi'_m)\ dS_{\xi};$$

here the integrations are carried out (not over the loci $r = \text{const.}$, $\xi_i = \text{const.}$

† For L_k an alternative procedure is to map the neighborhood of $r = 0$ on a concentric family of spheres in E_n, radii going into radii of the same length and angles between radii being preserved. The operation $L_k(u)$ for N_n goes into an operation L_k for E_n. As $L_k(u)$ is of class C'' at $r = 0$ for the Euclidean case, it will also be of class C'' for the non-Euclidean case, so that $\Delta_2 L_k(u)$ will exist and be continuous at $r = 0$. Likewise $\Delta_2 u$, continuous at $r = 0$ for N_n, will be continuous for $r = 0$ in E_n; hence $L_k(\Delta_2 u)$ will be continuous at $r = 0$ in E_n and therefore also in N_n.

A similar procedure could also be followed for $L_{k,n}$.

as for $L_{k,n-m}$ but) over $r = \text{const.}$, $\omega = \text{const.}$, that is, along "equidistant" m-dimensional surfaces, a distance r away from a flat N_m; h_k, h'_k are solutions of

$$\Delta_2 h = kh;$$

and dS_ξ is the $(m-1)$-dimensional content of the space $r = 0$. This operator $L^*_{k,m}$ may be shown to be permutable with Δ_2 under conditions analogous to those of Theorem IV.

If the space N_{n-m} of $L_{k,m}$ recedes to infinity in a proper manner, one is led to an element of arc of the form

$$ds^2 = d\xi^2 + e^{2\sigma\xi}(dx_1{}^2 + \cdots + dx^2{}_{n-1}),$$

the loci $\xi = \text{const.}$ being equidistant horospheres, and to operators

$$L^*_{k,m}(u) = h_k(x_1, x_m) \int \overset{\infty}{\underset{-\infty}{\cdots}} \int h'_k(x'_1, \cdots, x'_m) \times u(\xi; x'_1, \cdots, x'_m; x_{m+1}, \cdots, x_{n-1})\, dx'_1 \cdots dx'_m,$$

where h_k, h'_k are solutions of

$$(\partial^2/\partial x_1{}^2 + \cdots + \partial^2/\partial x_m{}^2)h = kh;$$

the left hand operator is the Beltrami operator along a horosphere $\xi = 0$; $x_{m+1} = \text{const.}, \cdots, x_{n-1} = \text{const.}$ Again the conditions under which one might prove the permutability of $L^*_{k,m}$ and Δ_2 are analogous to those of Theorem IV.

9. *Further extensions; iterated Laplacians.* Because of the very nature of the above theorems the functions u have been assumed to be of class C''. The theory of potential indicates in what direction the above results may be extended to less restricted functions. Thus, for $n = 3$, any function u of class C'' may be considered as the potential due to gravitating matter of volume density $\nabla^2 u/4\pi$ (plus a harmonic function). If one rephrases the various theorems above in terms of potentials and mass distributions one has a formulation which might be applied to such functions u as the potentials due to point, line, and surface distributions, using $\int (\partial u/\partial n)\, d\sum/4\pi$ as the mass inside a surface $\sum$.

Following out these ideas, one arrives at an extension of, say, Theorem II, of the following form: if u is a function, in general of class C', then the function

$$L_k(u) = h_k(\omega) \int h'_k(\omega')\, u(r, \omega')\, d\omega'$$

is the potential at r, ω obtained by taking the masses whose potential is the function u and replacing each mass initially at r, ω' by a surface distribution of mass over that sphere of density $h_k(\omega)h'_k(\omega')$ per total area of that sphere (or $h_k(\omega)h'_k(\omega')/4\pi$ per unit of the solid angle ω).

If one considers, on the other hand, more regular functions u, say, of class $C^{(2n)}$ it follows by successive applications of the various theorems that the above operators are permutable with iterations of the Laplacian. Thus, for example, if u is of class $C^{(6)}$, then the functions u, $\nabla^2 u$, $\nabla^4 u$ are all of class C''. Hence by applying Theorem II to them, we get

$$\begin{aligned} L_k[\nabla^2(\nabla^4 u)] &= \nabla^2[L_k(\nabla^4 u)], \\ L_k[\nabla^2(\nabla^2 u)] &= \nabla^2[L_k(\nabla^2 u)], \\ L_k[\nabla^2 u] &= \nabla^2[L_k(u)] \quad , \end{aligned}$$

and, transforming each right hand member by means of the equation which follows, arrive at

$$L_k(\nabla^6 u) = \nabla^6 [L_k(u)].$$

It will be noted that while the proof outlined does not prove that $L_k(u)$ is of class $C^{(6)}$, it justifies the three fold application of ∇^2 to this function and proves the final result to be a continuous function. That $L_k(u)$ is of class $C^{(6)}$ may be proved by the methods used in § 3.

Returning to the permutability of L_k and ∇^2, it is, of course, trivial to remark that the sum of two operators L_k with the same or different values of the degree k, result in an operator which is permutable with ∇^2; likewise, for the sum of an infinite number of such operators, barring questions of convergence. We shall now show that in a certain sense such a sum constitutes the most general linear functional operator L, permutable with ∇^2 and which replaces a function u over each of a concentric family of spheres by a function depending upon the value of u over that sphere only.

We shall suppose that the operator is of the form

$$L(u) = \int G(r, \omega, \omega')u(r, \omega')\,d\omega'$$

and that the function G may be expanded above in series of a complete set of surface spherical harmonics in both ω and ω'. Let $h_{k,1}, h_{k,2}, \cdots, h_{k,N_k}$ be a complete set of orthogonalized surface harmonics of degree k, normalized over the unit sphere. Expanding G, we get

$$L(u) = \sum_{k,i} \sum_{k',i'} G_{k,i;k',i'}(r)h_{k,i}(\omega) \int h_{k',i'}(\omega')u(r, \omega')\,d\omega',$$

assuming term by term integration permissible. Now choose for u the harmonic function $h_{k',i'}(\omega)r^{k'}$ for some fixed values of k', i'. All but one of the integrals vanish. As pointed out at the end of § 1, $L(u)$ must be harmonic if u is harmonic; hence we conclude that

$$\sum_{k,i} G_{k,i;k',i'}(r)r^{k'}h_{k,i}(\omega)$$

is harmonic; therefore the functions $G_{k,i;k',i'}(r)r^{k'}$ must reduce to a constant multiple of r^k:

$$G_{k,i;k',i'}(r) = C_{k,i;k',i'}r^{k-k'}.$$

It remains to show that the constants C vanish for $k \neq k'$.

To prove this choose $u = h_{k',i'}(\omega)r^{-k'-n+2}$. This function is harmonic for $r > 0$; hence $L(u)$ which reduces to

$$\sum_{k,i} C_{k,i;k',i'}r^{k-2k'-n+2}h_k(\omega),$$

must also be harmonic for $r > 0$. This can happen only if C vanishes, or if the exponent of r reduces to $-k-n+2$ or to k. Now the last case can only be realized if $n = 2$, $k' = 0$. The proposition is thus proved except for this special case; that this case forms no exception may now be shown by choosing $u = \log r$ and proceeding as above. In this way one proves that the operator L must reduce to a sum of the operators L_k.

GENERAL ELECTRIC COMPANY,
SCHENECTADY, N. Y.

ON MORSE'S DUALITY RELATIONS FOR MANIFOLDS.*

By Arthur B. Brown.†

1. *Introduction.* Morse has proved ‡ certain duality relations of importance in the calculus of variations. They refer to manifolds with regular boundaries and in one set of relations to immersed sub-complexes, with some analytic restrictions.§ We show that Morse's relations can be derived from certain relations of Lefschetz ¶ involving the homology characters of manifolds, of their closed subsets and of the residual spaces, and from the Poincaré duality theorem. Morse's relations involving immersed sub-complexes are also a corollary of relations of van Kampen,‖ who in addition has results involving intersection numbers, and treats general immersed complexes. By our different methods we obtain some results given neither by Morse nor van Kampen, principally Theorems 6 and 7.

We mention the possibility of extending the results for immersion to the case of an arbitrary closed sub-set of a manifold, for example by the method of infinite complexes, Lefschetz I, Chap. VII.

All results and proofs hold for combinatorial manifolds, and are valid if Betti numbers and homologies are taken either absolute or (mod. p), p any prime greater than unity.

For brevity, we shall generally understand "independent" to mean "linearly independent with respect to homologies".

2. *Manifolds with regular boundaries.*

* Presented to the Society, March 25, 1932.

† Part of the work on this paper was done while the author was a National Research Fellow.

‡ Marston Morse, "The Analysis and Analysis Situs of Regular n-Spreads in $(n+r)$-Space," *Proceedings of the National Academy of Sciences*, Vol. 13 (1927), pp. 813-817. These relations were brought to my attention by Morse, who suggested that they might possibly be proved by combinatorial processes.

§ We use the terminology of S. Lefschetz's "Topology," *Colloquium Series*, Vol. 12, New York, 1930. (Lefschetz I.)

¶ S. Lefschetz, "Manifolds with a Boundary and Their Transformations," *Transactions of the American Mathematical Society*, Vol. 29 (1927), pp. 429-462. For specific theorems, we shall refer to Lefschetz I.

‖ E. R. van Kampen, "Eine Verallgemeinerung der Alexanderschen Dualitätssatzes," *Amsterdam Proceedings*, Vol. 31 (1928), pp. 899-905. Here, as in Morse's paper, outlines of proofs are given.

THEOREM 1. *Given a connected combinatorial manifold M_n with non-vacuous regular boundary M_{n-1}, let*

$$\rho_i = R_i(M_n) + R_{n-i-1}(M_n) - R_i(M_{n-1}), \qquad (i < n), \tag{2.1}$$

$$\sigma_i = \sum_{k=0}^{i} (-1)^{i-k} \rho_k. \tag{2.2}$$

Then

$$0 \leqq \rho_i, \qquad 0 \leqq \sigma_i, \qquad 0 = \sigma_{n-1}. \tag{2.3}$$

Proof. We apply the definitions and formulas of Lefschetz I, §§ 38 and 39, Chap. III (pp. 149-150), replacing K and L by M_n and M_{n-1} respectively. First we note that $t_p = 0$ for all $p < n$, as follows from its definition and Theorem I on page 153 of Lefschetz I.* From (21) and (22) on page 150 we then obtain

$$R_i(M_n) - R_i(M_{n-1}) + s_i = r_i \qquad (i < n). \tag{2.4}$$

Now from (21) we have

$$s_i = R_{i+1}(M_n;\, M_{n-1}) - r_{i+1}. \tag{2.5}$$

From (9), page 142 of Lefschetz I, we have

$$R_{i+1}(M_n;\, M_{n-1}) = R_{n-i-1}(M_n - M_{n-1}). \tag{2.6}$$

By Theorem II, page 154,

$$R_{n-i-1}(M_n - M_{n-1}) = R_{n-i-1}(M_n). \tag{2.7}$$

Using (2.5), (2.6) and (2.7) we obtain

$$s_i = R_{n-i-1}(M_n) - r_{i+1}. \tag{2.8}$$

Substituting from (2.8) in (2.4) and using (2.1), we have

$$\rho_i = R_i(M_n) + R_{n-i-1}(M_n) - R_i(M_{n-1}) = r_i + r_{i+1}. \tag{2.9}$$

From the hypotheses of the theorem it is evident that $r_0 = r_n = 0$. Using that result together with (2.9), the relations (2.3) become obvious. In fact we have the following result.

THEOREM 2. *The difference in the first inequality of* (2.3) *is $r_i + r_{i+1}$; in the second is $r_i = r_i(M_n - M_{n-1})$.*

Theorems 1 and 2 hold for Betti numbers (mod. m), m any integer greater than unity, since the proofs are valid for that case. From the proofs

* Professor Lefschetz has called to my attention that throughout § 44 the words "regular or" should be deleted wherever they occur, as semi-regularity is required.

it is evident that *Theorems* 1 *and* 2 *are valid if* M_{n-1} *is not restricted to be a regular boundary, but is any sub-complex* L *such that* $M_n - L$ *is an open (combinatorial) manifold, and such that* $t_p = 0$ *for all* $p < n$.

3. *Residual complexes.*

THEOREM 3. *Given a connected combinatorial manifold* M_n, *and a sub-complex* K *such that neither* K *nor* $M_n - K$ *is vacuous, then if we set*

$$\rho_i = R_i(K) + R_{n-i-1}(M_n) - R_{n-i-1}(M_n - K) - \delta_0{}^i,$$

$$\sigma_i = \sum_{k=0}^{i} (-1)^{i-k} \rho_k,$$

*the following relations will hold.**

$$(3.1)\quad 0 \leqq \rho_i; \qquad (3.2)\quad 0 \leqq \sigma_i; \qquad (3.3)\quad 0 = \sigma_{n-1}.$$

Proof. First we take the case that K is a combinatorial n-manifold with regular boundary M_{n-1}. Then $B = M_n - K + M_{n-1}$ is likewise an n-manifold with regular boundary M_{n-1}.†

We introduce the following maximal sets of i-cycles, $i = 0, 1, \cdots, n$: A_i, on K, independent on K (as regards homologies) of i-cycles on M_{n-1}; B_i, on B, independent on B of i-cycles on M_{n-1}; $C_i{}^a$, on M_{n-1}, each bounding on K, all independent on B; $C_i{}^b$, on M_{n-1}, each bounding on B, all independent on K; $C_i{}^1$, on M_{n-1}, independent on M_n; $C_i{}^d$, on M_{n-1}, independent on M_{n-1}, each bounding both on K and on B. If we denote by a_i the number of i-cycles in $A_i, \cdots$, the numbers a_i, $b_i, \cdots$, are topological invariants. The following formulas hold.‡

$$(3.4)\qquad R_i(K) = a_i + c_i{}^1 + c_i{}^b.$$

$$(3.5)\qquad R_i(B) = b_i + c_i{}^1 + c_i{}^a.$$

$$(3.6)\qquad R_i(M_{n-1}) = c_i{}^1 + c_i{}^a + c_i{}^b + c_i{}^d.$$

$$(3.7)\qquad R_i(M_n) = a_i + b_i + c_i{}^1 + c^d{}_{i-1}.$$

Using (20), page 149, of Lefschetz I, we have:

$$(3.8)\qquad a_i = a_{n-i}.$$

$$(3.9)\qquad b_i = b_{n-i}.$$

* $\delta_0{}^i = 1$ or 0 according as $i = 0$ or $i \neq 0$.

† Lefschetz I, page 145, Corollary.

‡ A. B. Brown, *American Journal of Mathematics*, Vol. 52 (1930), pp. 251-270; Lemmas 1, 2, 4, 6. The formulas also suggest maximal independent sets of i-cycles for K, B, M_{n-1} and M_n, which are all actually proved to be valid. While the theorems are stated for the absolute or (mod. 2) case, all proofs are valid for the (mod. p) case as well.

Again using the relation $r_i = r_{n-i}$, we rewrite (2.8) in the form $R_i(M_n) - r_i = s_{n-i-1}$ and apply the result to the present K and M_{n-1} in place of M_n and M_{n-1} of § 2, thus obtaining (3.10) below; with B and M_{n-1} we obtain (3.11). We omit any proof beyond referring to (3.4), (3.5), (3.6) and the definitions above.

(3.10) $$c_i{}^b + c_i{}^1 = c^a{}_{n-i-1} + c^d{}_{n-i-1}.$$

(3.11) $$c_i{}^a + c_i{}^1 = c^b{}_{n-i-1} + c^d{}_{n-i-1}.$$

In the following, (3.12) follows from (3.6) and the Poincaré duality theorem for M_{n-1}; (3.13) from (3.10), (3.11) and (3.12); (3.14) from (3.11), (3.11) with i replaced by $n - i - 1$, and (3.13).

(3.12) $$c_i{}^1 + c_i{}^a + c_i{}^b + c_i{}^d = c^1{}_{n-i-1} + c^a{}_{n-i-1} + c^b{}_{n-i-1} + c^d{}_{n-i-1}.$$

(3.13) $$c_i{}^1 + c^1{}_{n-i-1} = c_i{}^d + c^d{}_{n-i-1}.$$

(3.14) $$c_i{}^a + c^a{}_{n-i-1} = c_i{}^b + c^b{}_{n-i-1}.$$

By a duality relation of Lefschetz we have (cf. (2.6))

(3.15) $$R_i(K - M_{n-1}) = R_{n-i}(K; M_{n-1}).$$

This gives us the relation

(3.16) $$a_i + c_i{}^1 + c_i{}^b = a_{n-i} + c^d{}_{n-i-1} + c^a{}_{n-i-1}.$$

That the left-hand sides of (3.15) and (3.16) are equal follows from (3.4) and the relation $R_i(K - M_{n-1}) = R_i(K)$. (Cf. (2.7)). The proof that the right-hand sides are equal is easily given by use of (21) on page 150 of Lefschetz I, using the fact that $t_i = 0$ (cf. § 2).

The following equations are derived as explained below.

(3.17) $$c_i{}^1 + c^d{}_{i-1} = c^1{}_{n-i} + c^d{}_{n-i-1}.$$

(3.18) $$c_i{}^d + c^d{}_{i-1} = c^1{}_{n-i} + c^1{}_{n-i-1}.$$

(3.19) $$c_0{}^d = c^1{}_{n-1}.$$

(3.20) $$c_i{}^1 = c^d{}_{n-i-1}, \qquad (i = 0, 1, \cdots, n-1).$$

(3.21) $$c_i{}^a = c^b{}_{n-i-1}, \qquad (i = 0, 1, \cdots, n-1).$$

Here (3.17) is obtained from the Poincaré duality relation for M_n, with the use of (3.7), (3.8) and (3.9); (3.18) from (3.13) and (3.17) by subtraction; (3.19) from (3.18) by setting $i = 0$ and observing that $c^d{}_{-1} = c_n{}^1 = 0$, as follows from the facts that there are no cycles of dimension

-1, and that M_{n-1} is of dimension only $n-1$; (3.20) from (3.18) and (3.19) by a succession of simple steps of algebra; (3.21) by substituting from (3.8) and (3.20) in (3.16), and replacing i by $n-i-1$.

We can now establish (3.1), (3.2) and (3.3). First we substitute from (3.21), with i replaced by $n-i-1$, in (3.4); then from (3.20) with i replaced by $n-i-1$, and from (3.8), in (3.7), to obtain the following relations.

$$R_i(K) = a_i + c^1_{n-i-1} + c_i^1. \tag{3.22}$$

$$R_i(M_n) = a_{n-i} + b_i + c_i^1 + c^1_{n-i}. \tag{3.23}$$

Since B is an n-manifold with regular boundary, $R_i(B) = R_i(M_n - K)$. We also have the obvious relations $a_0 = a_n = c_n^1 = 0$, $c_0^1 = 1$. By using these relations and (3.5), (3.22) and (3.23), we find upon substitution that (3.1) would be an equality if $a_i + a_{i+1} + c_i^1 + c^1_{i+1} - \delta_0^i$ were added to its left-hand side; that (3.2) would be an equality if $a_{i+1} + c^1_{i+1}$ were added to its left-hand side; and that (3.3) is valid. As $c_i^1 \geqq \delta_0^i$, Theorem 3 is therefore proved for the case that M_n is an n-manifold with regular boundary.

We consider now the case that K is any sub-complex of M_n. We denote by R' and R'' the first and second regular subdivisions, respectively, of any complex R. Let N' be the M'_n-neighborhood of K', less K'. (Thus $N' \not\supset K'$.) Let C be the complex which is the point-set boundary of the D''-neighborhood of K''. Then $N' + K'$ is a normal neighborhood of K', and N' is covered by a field F of curves, where each curve has its end-points on the boundaries on $N' + K'$ and of K' respectively, and cuts C in one point.* Hence N' is homeomorphic to the product P of C by a 1-cell. As N' is open on M'_n and M'_n is an n-manifold, N' is an open n-manifold; hence, by the topological invariance of the combinatorial manifold,† P is an open n-manifold. Therefore, if E_i is an i-cell of P, $LK_P(E_i)$ is an H_{n-i-1} (Lefschetz I, page 120, condition (b)). But if E'_{i-1} is the $(i-1)$-cell of C which generates E_i, then $LK_C(E'_{i-1})$ has the same structure as $LK_P(E_i)$, hence is likewise an $(n-i-1)$-sphere. Hence C, composed of simplicial cells, is an $(n-1)$-manifold. Consequently, if $K_1 = K'' + N'' + C$ and $B_1 = M_n'' - K_1$, then K_1 and $B_1 + C$ are n-manifolds each with regular boundary C.‡

From the proof for the case first considered it then follows that the relations of Theorem 3 are valid for A_1 and B_1 in place of K and $M_n - K$.

* Cf. Lefschetz I, page 91.

† E. R. van Kampen, "Die kombinatorische Topologie und die Dualitätssätze," *Leyden thesis* (1929), The Hague. Cf. Lefschetz I, page 155, § 4, Theorem.

‡ Lefschetz I, page 145, Corollary.

Now by use of the field F of curves in N' it is easily shown that K_1 and B_1 have the same Betti numbers as K and $M_n - K$, respectively.* As a consequence, the relations of Theorem 3 are valid for the given K, and the proof is complete.

We now state five theorems obtained (see below) from the following relations: Theorem 4 from the formulas for the differences in the sides of (3.1) and (3.2), mentioned at the end of the proof above for the case that K is a manifold with regular boundary; Theorem 5 from the same results, by symmetry; Theorems 6, 7, 8 from (3.20), (3.8) and (3.21) respectively. We first remark that Theorem 4 and the last part of Theorem 5 are not equivalent, since K is closed and $M_n - K$ is open, on M_n. The proofs of the theorems consist in demonstrating them, by use of the above-mentioned formulas, for the case that K is an n-manifold with regular boundary, and then extending the results to the case of a sub-complex, by use of C, N', etc., occurring at the end of the last proof. We shall omit the detailed proofs.

For brevity in stating the theorems, we introduce the following notation, where $R \supset L$: for the number of i-cycles in a maximal set —— on L, independent (with regard to homologies) on R, $\rho_i(L, R)$; on L, independent on R, but each $\sim$ an i-cycle on $R - L$, $\sigma_i(L, R)$; on R, independent of the i-cycles on L and those on $R - L$, $\tau_i(R, L)$; on L, independent on R of the i-cycles on $R - L$, $\mu_i(L, R)$; on L, independent on L, each bounding on R, $\nu_i(L, R)$.

THEOREM 4. *The difference between the two sides in* (3.1) *is* $\rho_i(K, M_n) + \rho_{i+1}(K, M_n) - \delta_0{}^i$, *and the difference for* (3.2) *is* $\rho_{i+1}(K, M_n)$.

THEOREM 5. *Under the hypotheses of Theorem* 3, *relations* (3.1), (3.2), (3.3) *remain valid if every* $R_j(K)$ *is replaced by* $R_j(M_n - K)$ *and every* $R_j(M_n - K)$ *by* $R_j(M_n)$. *The difference between the two sides in the relation corresponding to* (3.1) *is then* $\rho_i(M_n - K, M_n) + \rho_{i+1}(M_n - K, M_n) - \delta_0{}^i$, *and the difference for the relation corresponding to* (3.2) *is* $\rho_{i+1}(M_n - K, M_n)$.

THEOREM 6. *If* M_n *is a combinatorial* n*-manifold, and* K *a sub-complex, then* $\sigma_i(K, M_n) = \tau_{n-i}(M_n, K)$.

THEOREM 7. *Under the hypotheses of Theorem* 6,

$$\mu_i(K, M) = \mu_{n-i}(K, M_n).\dagger$$

* The result may also be obtained simply by use of the author's Lemma A, *Annals of Mathematics*, Vol. 32 (1931), p. 514.

† Theorems 6 and 7 together give van Kampen's relation 3°, for the present case, insofar as the number of cycles in each set is involved.

THEOREM 8. *Under the hypotheses of Theorem 6,*

$$\nu_i(K, M_n) = \nu_{n-i-1}(M_n - K, M_n).*$$

4. *Manifolds with the Betti numbers of an n-sphere.* If M_n has the Betti numbers of an n-sphere,† the quantities appearing in Theorems 4, 5, 7 become all zero, and those in Theorem 6 equal $\delta_0{}^i$. Hence Theorems 6 and 7 become trivial, and Theorems 3, 4, 5 and 8 reduce to the Alexander duality theorem for sub-complexes. But Theorem 1 gives an interesting result (of Morse), when K is immersed in ‡ M_n, as follows.

THEOREM 9. *Let K be a combinatorial n-manifold with non-vacuous regular boundary M_{n-1}, and M_n a combinatorial n-manifold with the Betti numbers of an n-sphere. If K can be immersed in M_n, then the following relations are valid:*

$$R_i(M_{n-1}) = R_i(K) + R_{n-i-1}(K). \tag{4.1}$$

Proof. Since (4.1) for K would be a consequence of (4.1) for the separate connected parts of K, obtained by summing, it will be sufficient to prove (4.1) for the case that K is connected. Then the hypotheses of Theorem 1 are satisfied, and, according to Theorem 2, (4.1) will be proved if we show that any i-cycle on K is $\approx$ on K to an i-cycle on M_{n-1}.

Since K is not in general a sub-complex of M_n, the result is not obvious. However, by using the data on the Betti numbers of M_n, and a field of curves in a normal neighborhood of M_{n-1} *on* K (obtained by regular subdivision; cf. the last part of the proof of Theorem 3), the proof is easily carried through, and we shall not give the details.

COLUMBIA UNIVERSITY.

---

* Theorem 8 corresponds to van Kampen's 1°, for our case, without, of course, giving his results regarding looping.

† This does not necessarily make M_n a combinatorial n-sphere, as no condition is imposed on the torsion coefficients of D.

‡ "Immersed in" means "coincident, as a result of a homeomorphism, with a sub-set of."

INVARIANTS OF INTERSECTION OF TWO CURVES ON A SURFACE.

By Ernest P. Lane.

1. *Introduction.* In a recent note * Bompiani has developed some stimulating ideas relative to certain invariants of intersection of two skew curves in ordinary space. He shows, in the first instance, that if two curves C, C' pass through a point P with distinct tangents t, t' at P and also with distinct osculating planes at P, whose line of intersection is different from t and from t', then there exist, through P and in the plane determined by t, t', two lines which are *principal* in the sense that the two cones projecting C, C' from any point on either line have contact of the second or higher order along their common generator through P, instead of contact of the first order as would ordinarily be the case if the center of projection were chosen at random in the plane of t, t'. Bompiani shows that the two principal lines separate the tangents t, t' harmonically. Moreover, he shows that on each principal line there is a point which is *principal* in the sense that if the projecting cones have their vertices at this point, the cones have contact of the third order.

In the case that the tangents t, t' are distinct but the osculating planes are coincident, Bompiani shows the existence of *three principal lines,* and of a principal point on each of these lines. He applies these considerations to the interesting case in which the two curves are asymptotic curves on a surface, and deduces some remarkable results which need not be reported in entirety here. But we shall refer to one of these results later on in Section 4.

The present note had its inception with the idea of applying the considerations summarized in the opening paragraph above to the case in which the two curves belong respectively to the two families of a conjugate net of curves on a surface. Some interesting results were found which will be presented forthwith. Moreover, some kindred considerations concerning certain quadric surfaces containing the asymptotic tangents at a point of a surface will be adjoined.

2. *Principal Lines.* We proceed to indicate an analytic basis for the

* Bompiani, "Invarianti d'intersezione di due curve sghembe," *Rendiconti dei Lincei,* Ser. 6, Vol. 14 (1931), pp. 456-461.

study of a surface referred to a conjugate net, and to determine the two principal lines, at a point of the surface, of the two curves of the net that pass through the point.

If the projective homogeneous coördinates $x^{(1)}, \cdots, x^{(4)}$ of a point P_x in ordinary space are given as analytic functions of two (and not fewer) independent variables u, v by equations of the form

$$x = x(u, v), \tag{1}$$

the locus of P_x as u, v vary is a proper analytic surface S. If the parametric curves on S form a conjugate net, the four coördinates x and the four coördinates y of a point on *the axis* of the point P_x satisfy a completely integrable system of linear partial differential equations of the form *

$$\begin{aligned} x_{uu} &= px + \alpha x_u + Ly, \\ x_{uv} &= cx + ax_u + bx_v, \\ x_{vv} &= qx + \delta x_v + Ny \qquad (LN \neq 0). \end{aligned} \tag{2}$$

The ray-points, or Laplace transformed points, x_{-1}, x_1 of the point P_x are given by the formulas

$$x_{-1} = x_u - bx, \qquad x_1 = x_v - ax. \tag{3}$$

The expansion for the u-curve at the point P_x of the form used by Bompiani in his note already cited may be calculated in the following way. The coördinates X of a point near P_x and on the u-curve through P_x are expressible by Taylor's expansion as power series in the increment Δu corresponding to displacement from P_x to the point X along the u-curve,

$$X = x + x_u \,\Delta u + x_{uu}\,\Delta u^2/2 + x_{uuu}\,\Delta u^3/6 + \cdots.$$

Expressing each of x_{uu}, x_{uuu} as a linear combination of x, x_{-1}, x_1, y we find

$$X = y_1 x + y_2 x_{-1} + y_3 x_1 + y_4 y,$$

where the local coördinates $y_1, \cdots, y_4$ of the point X are given by the expansions

$$\begin{aligned} y_1 &= 1 + b\,\Delta u + (p + b\alpha)\Delta u^2/2 + \cdots, \\ y_2 &= \Delta u + \alpha\,\Delta u^2/2 + \cdots, \\ y_3 &= \boldsymbol{H}\Delta u^3/6r + \cdots, \\ y_4 &= L\,\Delta u^2/2 + L[\alpha + b - (\log r)_u]\,\Delta u^3/6 + \cdots, \end{aligned} \tag{4}$$

and we have placed

* Lane, *Projective Differential Geometry of Curves and Surfaces*, University of Chicago Press (1932), p. 138.

$$(5)\qquad r = N/L, \qquad \boldsymbol{H} = b_v + ab + c - \delta_u.$$

Introducing non-homogeneous coördinates by the definitions

$$(6)\qquad x = y_2/y_1, \qquad y = y_3/y_1, \qquad z = y_4/y_1,$$

we find

$$(7)\qquad \begin{aligned} x &= \Delta u + (\alpha - 2b)\Delta u^2/2 + \cdots, \\ y &= \boldsymbol{H}\,\Delta u^3/6r + \cdots, \\ z &= L\,\Delta u^2/2 + L[\alpha - 2b - (\log r)_u]\,\Delta u^3/6 + \cdots. \end{aligned}$$

Inverting the first of these series we obtain

$$\Delta u = x - (\alpha - 2b)x^2/2 + \cdots,$$

and substituting this series for Δu in the last two of the series (7) we arrive at the required expansions for the u-curve, namely,

$$(8)\qquad y = \boldsymbol{H}x^3/6r + \cdots, \qquad z = Lx^2/2 + 4L\mathfrak{C}'x^3/3 + \cdots,$$

where we have placed

$$(9)\qquad 8\mathfrak{C}' = 4b - 2\alpha - (\log r)_u.$$

Similar calculations lead to the following expansions for the v-curve at the point P_x,

$$(10)\qquad x = r\boldsymbol{K}y^3/6 + \cdots, \qquad z = Ny^2/2 + 4N\mathfrak{B}'y^3/3 + \cdots,$$

where we have placed

$$(11)\qquad \boldsymbol{K} = a_u + ab + c - \alpha_v, \qquad 8\mathfrak{B}' = 4a - 2\delta + (\log r)_v.$$

Without further calculations we may compare the expansions (8), (10) with equations (1) on page 457 of Bompiani's note. Then making use of equations (5) on page 458 of Bompiani's note for the principal lines we reach immediately the following result:

The equations of the principal lines of the parametric curves at the point P_x of the surface S are

$$(12)\qquad z = Lx^2 - Ny^2 = 0.$$

The associate conjugate net of the parametric net on the surface S is by definition the unique conjugate net whose tangents (called *associate conjugate tangents*) at each point of S separate harmonically the parametric tangents at the point. The curvilinear differential equation of the associate conjugate net is

$$(13)\qquad L\,du^2 - N\,dv^2 = 0.$$

Consequently we have the theorem:

The principal lines, at a point of a conjugate net, of the two curves of the net that pass through the point are the associate conjugate tangents at the point.

3. *Principal Points and Principal Join.* We next determine the two principal points of the two curves of the parametric conjugate net at the point P_x, and study the line joining them.

Making use of equation (6) on page 458 of Bompiani's note we find that *the equation of the line joining the principal points is*

$$8(\mathfrak{C}'x + \mathfrak{B}'y) = 3. \tag{14}$$

We propose to call this line *the principal join* of the fundamental parametric curves at the point P_x. Solving equations (12) and (14) simultaneously one may obtain the non-homogeneous local coördinates of the principal points. Then passing from non-homogeneous local coördinates to general homogeneous coördinates one finds that *the points*

$$x_{-1} + 8\mathfrak{C}'x/3 \pm (x_1 + 8\mathfrak{B}'x/3)/r^{1/2} \tag{15}$$

are the principal points of the parametric curves at the point P_x.

Adding and subtracting the two formulas (15), one taken with the plus sign and one with the minus sign, we obtain the result:

The principal join crosses the parametric tangents at the point P_x in the points

$$x_{-1} + 8\mathfrak{C}'x/3, \qquad x_1 + 8\mathfrak{B}'x/3. \tag{16}$$

Recalling that a net is *quadratic*, i. e., lies on a quadric surface, in case

$$\mathfrak{B}' = \mathfrak{C}' = 0,$$

and further recalling that *the ray* of the parametric net crosses the parametric tangents at P_x in the ray-points x_{-1}, x_1, we have the theorem:

The principal join coincides with the ray at each point of a conjugate net if, and only if, the net is quadratic.

We shall suppose hereinafter that the fundamental conjugate net is not quadratic. It is not difficult to calculate the differential equation of the curves corresponding on the surface S to the developables of the congruence of principal joins, and to determine the focal surfaces of this congruence, but we shall have no immediate use for these results and so refrain from including them here.

The associate ray-points, i. e., the ray-points of the associate conjugate net, at the point P_x are easily shown to be given by the formulas

$$x_{-1} + 2\mathfrak{C}'x \pm (x_1 + 2\mathfrak{B}'x)/r^{1/2}. \tag{17}$$

Consequently *the associate ray*, which passes through these two points, crosses the parametric tangents at P_x in the points

$$x_{-1} + 2\mathfrak{C}'x, \qquad x_1 + 2\mathfrak{B}'x. \tag{18}$$

Taking suitable linear combinations of the formulas (16), (18) we may demonstrate the following theorem:

The ray, the associate ray, and the principal join at a point x of the parametric conjugate net are concurrent in the point z defined by

$$z = \mathfrak{B}'x_{-1} - \mathfrak{C}'x_1. \tag{19}$$

Moreover, a brief calculation, which we shall omit, suffices to demonstrate the following theorem:

The cross ratio of the ray, the associate ray, the line xz, and the principal join in the order named is 1/4.

4. *Studies in Asymptotic Parameters.* For many purposes, when studying conjugate nets on a surface S, it is convenient to take the asymptotic curves on the surface as parametric. In this section we shall make use of Fubini's canonical form of the differential equations of the surface, supposed not ruled and referred to its asymptotic curves, namely,

$$x_{uu} = px + \theta_u x_u + \beta x_v, \quad x_{vv} = qx + \gamma x_u + \theta_v x_v \quad (\theta = \log \beta\gamma). \tag{20}$$

The notation in this section will be that employed in Chapter III of the author's forthcoming book previously cited.

The curvilinear differential equation of any conjugate net N_λ on the surface S can now be written in the form

$$dv^2 - \lambda^2 du^2 = 0. \tag{21}$$

The equation of the associate conjugate net is obtained by changing the sign of λ^2.

By *a line* l_1 we mean any line through a general point P_x of the surface S and not lying in the tangent plane of S at P_x. Such a line joins the point x to the point y defined by placing

$$y = -ax_u - bx_v + x_{uv}, \tag{22}$$

where a, b are functions of u, v. Dually, by *a line* l_2 we mean any line in

the tangent plane of the surface S at the point P_x but not passing through P_x. Such a line joins the points ρ, σ defined by placing

$$\rho = x_u - bx, \qquad \sigma = x_v - ax, \tag{23}$$

where a, b are functions of u, v. When the functions a, b are the same in equations (23) as in (22), the lines l_1, l_2 are commonly called *reciprocal lines,* because they are reciprocal polar lines with respect to the quadric of Lie, whose equation referred to the tetrahedron x, x_u, x_v, x_{uv} with suitably chosen unit point is

$$2(x_2x_3 - x_1x_4) - (\beta\gamma + \theta_{uv})x_4^2 = 0. \tag{24}$$

It is known that *the ray* of the net N_λ is a line l_2, for which the functions a, b in (23) are given by

$$2a = \theta_v + (\log\lambda)_v - \beta/\lambda^2, \qquad 2b = \theta_u - (\log\lambda)_u - \gamma\lambda^2, \tag{25}$$

while for *the associate ray* it is only necessary to change the sign of λ^2. Moreover, it is known that *the flex-ray* of the net N_λ, i. e., the line of inflexions of the ray-point cubic of the pencil of conjugate nets determined by the net N_λ, is a line l_2 for which the functions a, b are given by

$$2a = \theta_v + (\log\lambda)_v, \qquad 2b = \theta_u - (\log\lambda)_u. \tag{26}$$

It is now easy to verify the conclusion:

The ray, associate ray, flex-ray, and principal join at the point P_x of the net N_λ are concurrent in the point z defined by placing

$$z = (\beta/\lambda^2)\{x_u - [\theta_u - (\log\lambda)_u]x/2\} - \gamma\lambda^2\{x_v - [\theta_v + (\log\lambda)_v]x/2\}. \tag{27}$$

In order to calculate *the functions a, b for the principal join* we make use of the concluding theorem of Section 3, and thus find

$$2a = \theta_v + (\log\lambda)_v + 5\beta/3\lambda^2, \qquad 2b = \theta_u - (\log\lambda)_u + 5\gamma\lambda^2/3. \tag{28}$$

By means of this result it is easy to verify the truth of the following statement:

The cross ratio of the ray, associate ray, flex-ray, and principal join in the order named at any point P_x of a net N_λ is $-1/4$.

The relation between a conjugate net and its associate conjugate net is obviously entirely reciprocal. So we are led to define *the associate principal join* at a point of a net N_λ, i. e., the principal join of the associate conjugate net of N_λ. For the associate principal join the functions a, b are obviously

obtained from the formulas (28) by merely changing the sign of λ^2. The associate principal join passes through the point z defined by (27). Moreover, it turns out that the cross ratio of the principal join, associate principal join, flex-ray, and line xz in the order named is -1. Thus we prove the theorem:

At a point x of a conjugate net N_λ, the principal join and associate principal join separate harmonically the flex-ray and the line xz connecting the point x to the point z of concurrency of the principal join and associate principal join.

We conclude these studies with some comments on certain non-singular quadric surfaces containing the asymptotic tangents at a point P_x of a surface S. The equation of any such quadric, referred to the tetrahedron x, x_u, x_v, x_{uv} with suitably chosen unit point, can be written in the form

$$x_2x_3 + x_4(k_1x_1 + k_2x_2 + k_3x_3 + k_4x_4) = 0, \tag{29}$$

where $k_1, \cdots, k_4$ are arbitrary. A line l_1 has the same polar line l_2 with respect to such a quadric as with respect to the quadric of Lie (24) if, and only if, the coefficients k_1, k_2, k_3 are connected by the relations

$$k_2 = b(1 + k_1), \qquad k_3 = a(1 + k_1). \tag{30}$$

Moreover, any quadric that has second-order contact with the surface S at the point P_x has an equation of the form (29) with $k_1 = -1$. Finally, the quadrics (29) for which $k_1 = -1$ and $k_2 = k_3 = 0$ are commonly called the quadrics of Darboux. Inspection of the relations (30) makes evident the following proposition:

If a quadric has contact of the second order at a point P_x of a surface S and if the polar relation of any two lines l_1, l_2 with respect to this quadric is equivalent to the polar relation of the lines with respect to the quadric of Lie at P_x, then the quadric is a quadric of Darboux of S at P_x.

It is further known that the quadrics (29) for which $k_1 = -3$ are the quadrics having contact of the third order with the asymptotic curves at the point P_x. Moreover, if also

$$k_2 = -\phi/2, \qquad k_3 = -\psi/2, \tag{31}$$

where

$$\phi = (\log \beta\gamma^2)_u, \qquad \psi = (\log \beta^2\gamma)_v, \tag{32}$$

it is known that a quadric (29) has contact of the fourth order with the asymptotic curves at P_x. Thus we see that *any quadric of the pencil*

$$x_2x_3 + x_4(-3x_1 - \phi x_2/2 - \psi x_3/2 + k_4x_4) = 0 \tag{33}$$

has contact of the fourth order with the asymptotic curves at the point P_x of the surface S. On page 461 of Bompiani's note previously cited he erroneously concludes that the pencil of quadrics having this property is the pencil of quadrics of Darboux. It is now proposed to name the pencil of quadrics (33) at the point P_x of the surface S *the principal pencil* of quadrics of S at P_x; any one of these quadrics may be called a *principal quadric.* If a unique covariant quadric of this pencil is desired, we may choose the one that passes through the covariant point $(0, 0, 0, 1)$. For this quadric we have $k_4 = 0$.

The polar line of the projective normal, joining the points $(1, 0, 0, 0)$, $(0, 0, 0, 1)$, with respect to any one of the principal quadrics (33) is the line l_2 for which

$$a = \psi/6, \qquad b = \phi/6. \tag{34}$$

The reciprocal polar line l_1 of this line l_2 with respect to the quadric of Lie is the canonical line called *the first principal line* at the point P_x. Thus a new characterization of this line is discovered, which should be compared with that given * by Fubini and Čech, and which may be stated in the following words:

The first principal line l_1 is the polar line with respect to any quadric of Darboux of that line l_2 which is the polar line of the projective normal with respect to any principal quadric.

Finally let us undertake to find a pair of lines l_1, l_2 which are reciprocal polars with respect to the quadric of Lie (or any quadric of Darboux) and are also reciprocal polars with respect to any principal quadric (33). Starting with any line l_1 we find its polar line with respect to any quadric of Darboux and also its polar line with respect to any quadric (33). Demanding that the latter two lines shall coincide we find that they coincide in the second canonical edge of Green, for which

$$a = \psi/4, \qquad b = \phi/4. \tag{35}$$

Thus we obtain a new characterization of the canonical edges of Green which may be formulated in the following theorem:

The canonical edges of Green are the only pair of lines l_1, l_2 which are reciprocal polars both with respect to the quadrics of Darboux and with respect to the principal quadrics.

UNIVERSITY OF CHICAGO,
CHICAGO, ILLINOIS.

---

* Fubini and Čech, *Geometria Proiettiva Differenziale*, Zanichelli, Vol. 1 (1926), p. 143.

SYSTEMS OF INVOLUTORIAL BIRATIONAL TRANSFORMATIONS CONTAINED MULTIPLY IN SPECIAL LINEAR LINE COMPLEXES.

By Evelyn Teresa Carroll.

Introduction. Birational transformations in space belonging to special line complexes have been treated by Pieri * and by H. A. Davis †; incidental mention of the subject has been made by various other authors. Pieri's method is synthetic, and in general no details have been given, while Davis has considered cases not treated here. The present paper contains a transformation not heretofore discussed and involves an interesting involution of fundamental elements along the basis line.

1. *Definition of the transformation.* Given a line $d \equiv x_1 = 0,\ x_2 = 0$ and a pencil of surfaces of order n, containing d to multiplicity $n-2$

$$\lambda_2 F_n(x) - \lambda_1 F'_n(x) = 0, \tag{1}$$

wherein

$$F_n(x) \equiv \sum_{\mu=0}^{n-2} x_1^{n-\mu-2} x_2^{\mu} u_\mu(x)$$

and

$$F'_n(x) \equiv \sum_{\mu=0}^{n-2} x_1^{n-\mu-2} x_2^{\mu} v_\mu(x),$$

$u_\mu(x)$ and $v_\mu(x)$ being quadratic forms in (x_1, x_2, x_3, x_4).

Let $(z) \equiv (0, 0, z_3, z_4)$ be a variable point on the line d, and let the pencil of surfaces (1) be connected with it by the relation

$$z_3\phi_1(\lambda_1, \lambda_2) - z_4\phi_2(\lambda_1, \lambda_2) = 0, \tag{2}$$

where ϕ_i $(i = 1, 2)$ is a binary form of order k.

A point (y) of space determines a surface of the pencil (1); so that (1) may be wrtten

$$F'_n(y)F_n(x) - F_n(y)F'_n(x) = 0 \tag{3}$$

and $z_3 = \phi_2[F_n(y), F'_n(y)]$ and $z_4 = \phi_1[F_n(y), F'_n(y)]$.

---

* M. Pieri, "Sulle trasformazioni birazionali dello spazio inerenti a un complesso lineare speciale," *Rendiconti del Circolo Matematico di Palermo*, Vol. 6 (1892), pp. 234-244.

† H. A. Davis, "Involutorial Transformations Belonging to a Linear Line Complex," *American Journal of Mathematics*, Vol. 52 (1930), pp. 53-71.

A point (x) on the line joining (y) to (z) has coördinates of the form

$$(4) \qquad \rho x_1 = \tau y_1, \quad \rho x_2 = \tau y_2, \quad \rho x_3 = \tau y_3 + \sigma\phi_2(F), \quad \rho x_4 = \tau y_4 + \sigma\phi_1(F).$$

The residual point (y') in which this line meets (3) again may be determined by applying the equations of transformation (4) to (3). The result in its simplified form is

$$\sigma[F'_n(y) \sum_{\mu=0}^{n-2} y_1^{n-\mu-2}y_2^{\mu}u_\mu(\phi) - F_n(y) \sum_{\mu=0}^{n-2} y_1^{n-\mu-2}y_2^{\mu}v_\mu(\phi)]$$
$$+ 2\tau[F'_n(y) \sum_{\mu=0}^{n-2} y_1^{n-\mu-2}y_2^{\mu}u_\mu(y, \phi) - F_n(y) \sum_{\mu=0}^{n-2} y_1^{n-\mu-2}y_2^{\mu}v_\mu(y, \phi)] = 0,$$

in which $u_\mu(\phi)$ and $v_\mu(\phi)$ are quadratic forms in

$$[0, 0, \phi_2(F_n(y), F'_n(y)), \phi_1(F_n(y), F'_n(y))]$$

and $u_\mu(y, \phi)$ and $v_\mu(y, \phi)$ are the polar forms of

$$(z) \equiv [0, 0, \phi_2(F_n(y), F'_n(y)), \phi_1(F_n(y), F'_n(y))]$$

in regard to $u_\mu(y)$ and $v_\mu(y)$, respectively. Therefore

$$\sigma \equiv -2[F'_n(y) \sum_{\mu=0}^{n-2} y_1^{n-\mu-2}y_2^{\mu}u_\mu(y, \phi) - F_n(y) \sum_{\mu=0}^{n-2} y_1^{n-\mu-2}y_2^{\mu}v_\mu(y, \phi)]$$
$$\tau \equiv F'_n(y) \sum_{\mu=0}^{n-2} y_1^{n-\mu-2}y_2^{\mu}u_\mu(\phi) - F_n(y) \sum_{\mu=0}^{n-2} y_1^{n-\mu-2}y_2^{\mu}v_\mu(\phi).$$

The transformation is of order $2n(k+1)-1$, where τ and σ have the orders $2[n(k+1)-1]$ and $n(k+2)-1$, respectively. The line d appears to multiplicity $(n-2)(k+2)$ on $\sigma = 0$ and to multiplicity $2(n-2)(k+1)$ on $\tau = 0$ and on the surfaces which are conjugates of planes in the transformation.

The image of the line d is the surface $\tau = 0$; $\sigma = 0$ is the locus of invariant points.

2. *Fundamental basis curve.* The surfaces $F_n(x) = 0$ and $F'_n(x) = 0$ intersect in a curve of order n^2, composed of d to multiplicity $n-2$ on each surface and a curve $\gamma_{4(n-1)}$ of order $4(n-1)$. The latter is $k+1$ fold on $\sigma = 0$; it is $2k+1$ fold on $\tau = 0$, and on the surfaces conjugate to the planes of space.

If a plane is passed through d, it will intersect $F_n(x)$ in a curve of order n, which consists of d taken $(n-2)$ times and a conic. This is likewise true for $F'_n(k)$ and the two conics intersect in four points. As a plane intersects $\gamma_{4(n-1)}$ in $4(n-1)$ points, $\gamma_{4(n-1)}$ and d must intersect in $4(n-2)$ points.

3. *Plane transformation in planes through d.* Since every plane through d is transformed into itself, it is expedient to discuss the plane transformation. The plane meets $\gamma_{4(n-1)}$ in four points, A, B, C, and D, the basis of a pencil of conics.

In the plane this pencil of conics may be represented by the equation

$$\lambda_2 C_1(x) - \lambda_1 C_2(x) = 0, \tag{5}$$

where $C_1(x) \equiv \sum_1^3 a_{ij}x_ix_j$ and $C_2(x) \equiv \sum_1^3 b_{ij}x_ix_j$, and the line d may be considered to have the equation $x_3 = 0$. Let the pencil of conics and the point $(z) \equiv (z_1, z_2, 0)$ be connected by the relation

$$z_1\phi_2(\lambda_1, \lambda_2) - z_2\phi_1(\lambda_1, \lambda_2) = 0, \tag{6}$$

where ϕ_i $(i = 1, 2)$ is a binary form of order k.

A point (y) in the plane determines a conic of the pencil; hence

$$C_2(y)C_1(x) - C_1(y)C_2(x) = 0, \tag{7}$$

and

$$z_1 = \phi_1[C_1(y), C_2(y)], \qquad z_2 = \phi_2[C_1(y), C_2(y)]. \tag{8}$$

A point (x) on the line joining (y) to (z) has coördinates of the form

$$x_1 = \tau y_1 + \sigma\phi_1(C), \qquad x_2 = \tau y_2 + \sigma\phi_2(C), \qquad x_3 = \tau y_3. \tag{9}$$

When these equations of transformation have been applied to (7) and the necessary expansion has been performed, the resulting form of the equation is

$$\sigma[C_1(\phi)C_2(y) - C_2(\phi)C_1(y)] + 2\tau[C_1(\phi, y)C_2(y) - C_2(\phi, y)C_1(y)] = 0$$

in which $C_i(\phi)$ $(i = 1, 2)$ are quadratic forms in $(\phi_1, \phi_2, 0)$ and $C_i(\phi, y)$ are the polar forms of $(\phi_1, \phi_2, 0)$ with respect to $C_i(y)$. Therefore

$$\sigma = -2[C_1(\phi, y)C_2(y) - C_2(\phi, y)C_1(y)]$$
$$\tau = C_1(\phi)C_2(y) - C_2(\phi)C_1(y).$$

The order of this transformation is $4k + 3$, while σ and τ have orders $2k + 3$ and $2(2k + 1)$, respectively. The points A, B, C, and D occur to multiplicity $2k + 1$ on the curve $\tau = 0$ and on the general curves, conjugates of the straight lines of the plane; and to multiplicity $k + 1$ on the curve $\sigma = 0$, which is the locus of invariant points.

The class of the involution is k, as any line contains k pairs of conjugates; a straight line is transformed into a curve of order $4k + 3$, which meets the

given line in $4k+3$ points, $2k+3$ of which are fixed, the remaining $2k$ points being arranged in k pairs of conjugates.

$\tau=0$ is composite, being composed of $2k+1$ conics of the pencil. Every point in which any one of these $2k+1$ conics meets $\sigma=0$ is a fundamental point. The number of such intersections is $4k+6$, $4(k+1)$ of which fall at the base points; this leaves two points of each conic on $\sigma=0$ and they may be distinct, consecutive, or coincident. No additional fundamental points exist as can be seen from the equations of the involution.

Since the involution is a Cremona transformation, the two fundamental equations $\Sigma\,\alpha_i\,i=3(n-1)$ and $\Sigma\,\alpha_i\,i^2=n^2-1$ must be satisfied. Here $n=4k+3$ and $\alpha_{2k+1}=4$; either $\alpha_1=4k+2$ or $\alpha_2=2k+1$ and all other α_i's equal to zero satisfy the first equation; the second alternative satisfies the second equation.

Hence *

$$C_{4k+3}:\ 4^{2k+1}(2k+1)^2.$$

There are $2k+1$ fundamental conics belonging to the given pencil and each passing through one of the double points, and four curves of order $2k+1$, images of A, B, C, and D, respectively; these latter curves have one of the base points to multiplicity $k+1$ and each of the others to multiplicity k, while they pass simply through the $2k+1$ fundamental double points on d.

The image of any point on $\tau=0$ is a definite point on the line $x_3=0$.

There are $2k+1$ conics of the system that pass through their associated points (z) on $x_3=0$; these are fundamental and the point (z) is transformed into itself, corresponding to the tangent to the associated conic passing through it. Hence the curve $\sigma=0$ passes through each of these points. There are also two conics of the pencil which are tangent to $x_3=0$; their points of contact belong to $\sigma=0$. These two categories account for all the intersections of $\sigma=0$ and $x_3=0$.

The line $x_3=0$ is transformed into itself but not point for point.

The curve $\sigma=0$ touches each conic of $\tau=0$ at one of the intersections of $\sigma=0$ and $x_3=0$.

The complete configuration of fundamental points consists of the four base points, each to multiplicity $k+1$, and the $2k+1$ double points on $x_3=0$. The image of each of the latter is the conic of the pencil $A\,B\,C\,D$ passing through it.

The image of a base point, as A, is a curve of order $2k+1$, having A

* Ruffini, "Sulla risoluzione delle due equazioni di condizione delle trasformazioni Cremoniane piane delle figure piane," *Atti dell'Accademia di Bologna, Memorie* (3), Vol. 8 (1877), pp. 456-525.

to multiplicity $k+1$ and passing k times through the other three base points and doubly through the $2k+1$ fundamental points on $x_3=0$. As (z) describes $x_3=0$, k conics of the pencil are uniquely determined by the tangent t at A; t and the associated line $A(z)$ are in $(1,k)$ correspondence and hence have $k+1$ self-corresponding elements in two concentric pencils.

The table of characteristics of this plane transformation may be expressed in the form

$$\begin{aligned}
&C_1 \sim C_{4k+3} : A^{2k+1}\, B^{2k+1}\, C^{2k+1}\, D^{2k+1}\, P_1{}^2 P_2{}^2 \cdots P^2{}_{2k+1} \\
&A \sim \alpha_{2k+1} : A^{k+1}\, B^k\, C^k\, D^k\, P_1 P_2 \cdots P_{2k+1} \\
&P_i \sim \pi_{2,i} : A\, B\, C\, D\, P_i \\
&J_{12k+6}(C_{4k+3}) : \alpha_{2k+1}\, \beta_{2k+1}\, \gamma_{2k+1}\, \delta_{2k+1} (2k+1)\,(\pi_{2,i})^2 \\
&\tau_{2(2k+1)} : (2k+1)\pi_{2,i} \\
&\sigma_{2k+3} : A^{k+1}\, B^{k+1}\, C^{k+1}\, D^{k+1}\, P_1 P_2 \cdots P_{2k+1} (\text{genus } k+1).
\end{aligned}$$

4. *Image of points of* $\gamma_{4(n-1)}$. The image of a point on $\gamma_{4(n-1)}$ is the plane curve of order $2k+1$ just obtained in the $(1,k)$ correspondence of the points of d and a pencil of conics.

A plane meets its own image surface in a composite curve of order $2n(k+1)-1$, consisting of a curve of order $n(k+2)-1$, the intersection of $\sigma=0$ and the given plane; and k curves of order n, the intersections of the plane with the k surfaces of order n belonging to the point (z) of intersection of the given plane and d. Each of these curves of order n goes into itself by central projection, center on d.

5. *Conjugate of a line.* The image of an arbitrary straight line l is a curve $C_{2n(k+1)-1}$ of order $2n(k+1)-1$ which meets d in $2[n(k+1)-1]$ points, the images of the intersections of l and $\tau_{2[n(k+1)-1]}=0$. This $C_{2n(k+1)-1}$ also meets l in $n(k+2)-1$ points, the intersections of l and $\sigma_{n(k+2)-1}=0$. If l meets d, the image of l loses a curve of order $2(n-2)\times(k+1)$, the image of the point of intersection of d and l, as d is $2(n-2)(k+1)$ fold in the transformation. The proper conjugate is a curve of order $4k+3$, lying in the plane determined by d and l; it has four $(2k+1)$ fold points at the intersections of the plane and $\gamma_{4(n-1)}$ and $2k+1$ double points on d.

Every plane through d cuts $\tau_{2[n(k+1)-1]}=0$ in d counted $2(n-2)(k+1)$ times and in $2k+1$ conics, the images of the $2k+1$ fundamental points on d in that plane.

6. *Image of* $\gamma_{4(n-1)}$. The image of a surface of order n of the pencil is of order $n[2n(k+1)-1]$ and is composed of the given surface, of $\tau_{2[n(k+1)-1]}$ to multiplicity $(n-2)$, and of a surface $\Gamma_{4[n(k+1)-1]}$ of order

$4[n(k+1)-1]$, the image of $\gamma_{4(n-1)}$. $\Gamma_{4[n(k+1)-1]}$ contains $\gamma_{4(n-1)}$ to multiplicity $4k+1$ and d itself to multiplicity $4(n-2)(k+1)$. This can be seen from the fact that a plane through d cuts $\Gamma_{4[n(k+1)-1]}$ in a curve of order $4[n(k+1)-1]$; d intersects $\gamma_{4(n-1)}$ in $4(n-2)$ points; each point of $\gamma_{4(n-1)}$ goes into a plane curve of order $2k+1$; therefore d appears on $\Gamma_{4[n(k+1)-1]}$ to multiplicity $4(n-2)(k+1)$. $[4(nk+n-1)-4(2k+1) = 4(n-2)(k+1)]$.

7. *Points on d.* The line d is $2(n-2)(k+1)$ fold in the transformation, hence any point on d has an image curve of order $2(n-2)(k+1)$. But any point of d has an image conic in each of the $k(n-2)$ tangent planes of the k surfaces belonging to that point, that is, a curve of order $2k(n-2)$ is accounted for; the residual is d itself taken $2(n-2)$ times. Hence d is a fundamental line of the first kind and also of the second for $n>2$. The equations of transformation show that any point on d goes into the whole line d.

When $n=2$, the line d is not a fundamental line. The resulting transformation in this case is described by Montesano.* In this case consider the base curve γ_4 of a pencil of quadrics and g a bisecant of γ_4. A point (y) on g determines the quadric H_g of the pencil containing g, and also determines the associated point (z) on d. The plane determined by (z) and g meets $H_g=0$ in a residual line g' through the two residual points of γ_4 in the plane. Thus, each quadric of the pencil is transformed into itself, and the two systems of reguli are interchanged. The $2k+1$ pairs of parasitic lines meet at $P_1, \cdots P_{2k+1}$ on d, hence these lines of each pair are also of different reguli.

8. *τ and σ tangent.* The surfaces $\tau_{2[n(k+1)-1]}=0$ and $\sigma_{n(k+2)-1}=0$ may be shown to be tangent along d, so that d counts for $2k$ units more in the common intersection. A plane passed through d intersects $\sigma_{n(k+2)-1}=0$ in a curve $C_{n(k+2)-1}$ of order $n(k+2)-1$, in which d is counted $(n-2)(k+2)$ times, leaving a curve C_{2k+3} of order $2k+3$. In the plane C_{2k+3} has $k+1$-fold points at each of A, B, C, and D, the intersections of the two conics cut from the base surfaces by the plane; and simple points in $P_1, P_2, \cdots, P_{2k+1}$, the $2k+1$ fundamental points on d in that plane. The plane intersects $\tau_{2[n(k+1)-1]}=0$ in $2k+1$ conics of the same pencil $ABCD$, which pass through $P_1, P_2 \cdots P_{2k+1}$, respectively. Each of these conics meets C_{2k+3} in $2(2k+3)$ points, consisting of $(k+1)$ each at A, B, C, and D and two

* Montesano, "Su una classe di trasformazioni razionali ed involutorie dello spazio di genere arbitrario n e di grado $2n+1$," *Giornale Matematiche di Battaglini*, Vol. 31 (1893), pp. 36-50.

at a fundamental point P_i. Every point of the conic goes into P_i and no point of the conic lies on C_{2k+8} except the fundamental points; hence the section of $\sigma_{n(k+2)-1}=0$ must touch that conic at P_i; the surfaces $\tau_{2[n(k+1)-1]}=0$ and $\sigma_{n(k+2)-1}=0$ are tangent along d. In fact, d counts for $2k$ units more in the common intersection.

At any point (z) on d $2k(n-2)$ tangent planes to a surface of the pencil coincide in pairs and with the $k(n-2)$ tangent planes to the k surfaces of order n which are associated with that point. These $k(n-2)$ planes are also tangent to $\tau_{2[n(k+1)-1]}=0$ and to $\sigma_{n(k+2)-1}=0$, each simply. The fundamental conics in these planes lie on all the surfaces of the web $|S_{2n(k+1)-1}|$, conjugate to the field of planes of space.

The other planes associated with the same point and tangent to $S_{2n(k+1)-1}$ form a group of an involution of order $2(n-2)$

$$z_3[2(n-2) \text{ order in } x_1, x_2] + z_4[2(n-2) \text{ order in } x_1, x_2] = 0;$$

when (z) describes d, the group of planes describe the involution.

9. *Analytic method.* The details of this analysis will be given for $n=3$, $k=1$, since this case can be seen more readily than the general case.

The tangent plane to the cubic surface

$$z_4(x_1u + x_2v) - z_3(x_1w + x_2t) = 0,$$

in which u, v, w, and t are quadratic forms in (x_1, x_2, x_3, x_4), at the point $(0, 0, z_3, z_4)$ is

$$z_4[x_1u(z) + x_2v(z)] - z_3[x_1w(z) + x_2t(z)] = 0.$$

A plane $(ax) = 0 \sim \tau(ax) + \sigma[a_3(x_1u + x_2v) + a_4(x_1w + x_2t)] = 0$. Call this image surface S_{11}. The configuration of the tangent planes to $\sigma_8 = 0$, $\tau_{10} = 0$, and $S_{11} = 0$ is found from those terms in their equations which contain the highest powers of x_3, x_4. These appear to degrees five, six, and seven, respectively, in σ_8, τ_{10}, and S_{11}. When z_3 and z_4 are substituted for x_3 and x_4, respectively, in these terms of σ_8, τ_{10}, and S_{11}, the results denoted by σ_0, τ_0, and S_0, respectively, have the forms

$$\begin{aligned}
\sigma_0 \equiv & -2[z_4(x_1u(z) + x_2v(z)) - z_3(x_1w(z) + x_2t(z))] \\
& \times [Az_3^2 + Bz_3z_4 + Cz_4^2] = 0, \\
\tau_0 \equiv & [z_4(x_1u(z) + x_2v(z)) - z_3(x_1w(z) + x_2t(z))] \\
& \times [2Az_3(x_1u(z) + x_2v(z) + B\{z_4(x_1u(z) + x_2v(z) \\
& \quad + z_3(x_1w(z) + x_2t(z))\} + 2Cz_4(x_1w(z) + x_2t(z))] = 0, \\
S_0 \equiv & [z_4(x_1u(z) + x_2v(z)) - z_3(x_1w(z) + x_2t(z))]^2 \\
& \times [a_4(2Az_3 + Bz_4) - a_3(Bz_3 + 2Cz_4)] = 0,
\end{aligned}$$

where

$$
\begin{aligned}
A &\equiv (x_1w_{34} + x_2t_{34})(x_1u_{33} + x_2v_{33}) - (x_1u_{34} + x_2v_{34})(x_1w_{33} + x_2t_{33}) \\
B &\equiv (x_1w_{44} + x_2t_{44})(x_1u_{33} + x_2v_{33}) - (x_1u_{44} + x_2v_{44})(x_1w_{33} + x_2t_{33}) \\
C &\equiv (x_1w_{44} + x_2t_{44})(x_1u_{34} + x_2v_{34}) - (x_1u_{44} + x_2v_{44})(x_1w_{34} + x_2t_{34}).
\end{aligned}
$$

Hence from the above forms it is seen that of the four tangent planes to $S_{11} = 0$ at any point $(0, 0, z_3, z_4)$ on d, two coincide with each other and with the tangent plane to the cubic surface $z_4(x_1u + x_2v) - z_3(x_1w + x_2t) = 0$ associated with that point. This plane is also tangent to $\tau_{10} = 0$ and $\sigma_8 = 0$, each simply. The fundamental conic in this plane lies on all the $| S_{11} |$ of the ∞^2 system, conjugate to the bundle of planes passing through the associated point on d.

The equations of the other two planes associated with the same point and tangent to $S_{11} = 0$ contain the coördinates of (z) simply; $a_3(Bz_3 + 2Cz_4) - a_4(2Az_3 + Bz_4) = 0$; hence when (z) describes d, the pair of planes describes an involution of order two.

Thus in the general case, the image of any point of d is a curve of order $2(n-2)(k+1)$, consisting of $k(n-2)$ conics and of the line d taken $2(n-2)$ times.*

Any two $S_{2n(k+1)-1}$ of the web $| S_{2n(k+1)-1} |$ intersect in d to multiplicity $4(n-2)^2[(k+1)^2+1]$. Every surface of the system has $(n-2)$ consecutive double lines d; the $k(n-2)$ tangent planes each counted twice being those of the associated surface of order n and being such double tangent planes for every $S_{2n(k+1)-1}$ of the system.

10. *Parasitic lines.* There are $[(6n-8)(k+1)-2]$ parasitic lines or fundamental straight lines of the second kind occurring simply on $\sigma_{n(k+2)-1} = 0$, $\tau_{2[n(k+1)-1]} = 0$, and hence on every $S_{2n(k+1)-1}$ of the web; they appear doubly on $\Gamma_{4[n(k+1)-1]} = 0$. The proof for determining the number of these parasitic lines may be given as follows:

The equation of the $n-2$ tangent planes to a surface

$$\lambda_2 \sum_{\mu=0}^{n-2} x_1^{n-\mu-2}x_2^{\mu}u_\mu(x) - \lambda_1 \sum_{\mu=0}^{n-2} x_1^{n-\mu-2}x_2^{\mu}v_\mu(x) = 0$$

of the pencil (1) at the point $(0, 0, z_3, z_4) \equiv (0, 0, \phi_2(\lambda), \phi_1(\lambda))$ is

$$\lambda_2 \sum_{\mu=0}^{n-2} x_1^{n-\mu-2}x_2^{\mu}u_\mu(\phi) - \lambda_1 \sum_{\mu=0}^{n-2} x_1^{n-\mu-2}x_2^{\mu}v_\mu(\phi) = 0. \tag{10}$$

Let one of these planes be expressed in the form $x_2 = rx_1$. When this sub-

* Montesano, "Sulla teoria generale delle corrispondenze birazionali dello spazio," *Rendiconti della R. Accademia dei Lincei*, Vol. 5 (1918), pp. 396-400, 438-441.

stitution is made in the equation of the surface (1), and the necessary reduction is performed, the result is

$$\lambda_2 \sum_{\mu=0}^{n-2} r^{\mu} u_{\mu}(x) - \lambda_1 \sum_{\mu=0}^{n-2} r^{\mu} v_{\mu}(x) = 0. \tag{11}$$

Let (11) be represented by $C_2 \equiv \sum_{i,k=1}^{3} a_{ik} x_i x_k = 0$. It is seen that a_{11} is a coefficient of order n in r and one in λ; a_{13} and a_{14} of order $n-1$ in r and one in λ; and a_{33}, a_{44}, and a_{34} of order $n-2$ in r and one in λ. The discriminant has highest powers of r, λ, thus:

$$\Delta \equiv \begin{vmatrix} r^{n}, \lambda & r^{n-1}, \lambda & r^{n-1}, \lambda \\ r^{n-1}, \lambda & r^{n-2}, \lambda & r^{n-2}, \lambda \\ r^{n-1}, \lambda & r^{n-2}, \lambda & r^{n-2}, \lambda \end{vmatrix},$$

hence it contains r to power $3n-4$ and λ to power three.

The tangent planes (10) contain r to power $n-2$ and λ to power $2k+1$, since $\phi(\lambda)$ is of order $2k$ in λ.

When λ is eliminated between the equation of the tangent planes (10)

$$\lambda_2 \sum_{\mu=0}^{n-2} r^{\mu} u(\phi) - \lambda_1 \sum_{\mu=0}^{n-2} r^{\mu} v_{\mu}(\phi) \equiv A_0 \lambda^{2k+1} + A_1 \lambda^{2k} + \cdots + A_{2k+1} = 0$$

and

$$\Delta \equiv B_0 \lambda^3 + B_1 \lambda^2 + B_2 \lambda + B_3 = 0,$$

the resultant is of the form $\Sigma A_i^3 \cdot B_j^{2k+1}$, where A_i is of order $n-2$ in r and B_j is of order $3n-4$ in r. Therefore there are

$$3(n-2) + (2k+1)(3n-4) = [(6n-8)(k+1)-2]$$

parasitic lines.

The plane of any one of these lines g and d is tangent to its associated surface $F_n(x) = 0$ at $(g, d) = (z)$. The residual conic cut from the surface by this plane is composite, one component g passing through (z). Any point of g goes into the whole line g; the other component is the image of the point (z).

11. *Table of characteristics.* The general table of characteristics has the form

$$S_1 \sim S_{2n(k+1)-1} : d^{2(n-2)(k+1)} k(n-2) \bar{d}^2 \gamma^{2k+1}_{4(n-1)} [(6n-8)(k+1)-2] g,$$
$$d \sim \tau_{2[n(k+1)-1]} : d^{2(n-2)(k+1)} k(n-2) \bar{d} \gamma^{2k+1}_{4(n-1)} [(6n-8)(k+1)-2] g,$$
$$\sigma_{n(k+2)-1} \sim \sigma_{n(k+2)-1} : d^{(n-2)(k+2)} k(n-2) \bar{d} \gamma^{k+1}_{4(n-1)} [(6n-8)(k+1)-2] g,$$
$$\gamma_{4(n-1)} \sim \Gamma_{4[n(k+1)-1]} : d^{4(n-2)(k+1)} k(n-2) \bar{d}^4 \gamma^{4k+1}_{4(n-1)} [(6n-8)(k+1)-2] g^2,$$
$$J_{8[n(k+1)-1]} : \quad \Gamma_{4[n(k+1)-1]} \quad \tau^2_{2[n(k+1)-1]}.$$

12. *Intersections of principal surfaces.* A detailed account of the intersections of the various principal elements with each other in the transformation may now be explained.

Two $S_{2n(k+1)-1}$ of the web intersect in a curve of order $2n(k+1)-1$, the image of a straight line which is the intersection of the two planes, the conjugates of the two given surfaces; d to multiplicity $[2(n-2)(k+1)]^2$; $\gamma_{4(n-1)}$ to multiplicity $(2k+1)^2$; the $(6n-8)(k+1)-2$ parasitic lines; and a consecutive double line in each of the k sheets of the two surfaces.

An $S_{2n(k+1)-1}$ of the web and $\tau_{2[n(k+1)-1]}$ intersect in d to multiplicity $4(n-2)^2(k+1)^2$; $\gamma_{4(n-1)}$ to multiplicity $(2k+1)^2$; the $(6n-8)\times(k+1)-2$ parasitic lines; and $k(n-2)$ conics and d counted $2k(n-2)$ times, the images of a point on d which is the intersection of a plane and d, the conjugates of the given surfaces and $\tau_{2[n(k+1)-1]}$, respectively.

An $S_{2n(k+1)-1}$ of the web and $\sigma_{n(k+2)-1}$ intersect in d to multiplicity $2(n-2)^2(k+1)(k+2)$; $\gamma_{4(n-1)}$ to multiplicity $(2k+1)(k+1)$; $(6n-8)(k+1)-2$ parasitic lines; a curve of order $n(k+2)-1$, the intersection of a plane and $\sigma_{n(k+2)-1}$, the conjugates of the given surfaces; and d which is double on the k sheets of $S_{2n(k+1)-1}$.

An $S_{2n(k+1)-1}$ of the web and $\Gamma_{4[n(k+1)-1]}$ intersect in d to multiplicity $4(n-2)^2(k+1)^2$; $\gamma_{4(n-1)}$ to multiplicity $(2k+1)(4k+1)$; the $(6n-8)\times(k+1)-2$ parasitic lines twice; $4(n-1)$ curves of order $2k+1$, the images of the points of intersection of a plane and $\gamma_{4(n-1)}$, which are conjugates of the given surfaces; and d taken $8k(n-2)$ times, as d is a consecutive double line on $S_{2n(k+1)-1}$ and a consecutive four-fold line on $\Gamma_{4[n(k+1)-1]}$.

$\tau_{2[n(k+1)-1]}$ and $\sigma_{n(k+2)-1}$ intersect in d to multiplicity $2(n-2)^2(k+1)\times(k+2)$; $\gamma_{4(n-1)}$ to multiplicity $(2k+1)(k+1)$; the $[(6n-8)\times(k+1)-2]$ parasitic lines; and d taken $2k(n-2)$ times, since d counts for $2k$ units more in the common intersection.

$\tau_{2[n(k+1)-1]}$ and $\Gamma_{4[n(k+1)-1]}$ intersect in d to multiplicity $8(n-2)^2\times(k+1)^2$; $\gamma_{4(n-1)}$ to multiplicity $(2k+1)(4k+1)$; the $(6n-8)\times(k+1)-2$ parasitic lines twice; $4k(n-2)$ conics in the four tangent planes of the k surfaces associated with the points in which $\gamma_{4(n-1)}$ meets d; and d taken $4k(n-2)$ times, since d is four-fold on k sheets of $\Gamma_{4[n(k+1)-1]}$
$\times(k+1)$; $\gamma_{4(n-1)}$ to multiplicity $(k+1)(4k+1)$; $(6n-8)(k+1)-2$

$\sigma_{n(k+2)-1}$ and $\Gamma_{4[n(k+1)-1]}$ intersect in d to multiplicity $4(n-2)^2(k+2)$ and single on $\tau_{2[n(k+1)-1]}$.

parasitic lines twice; d taken $4k(n-2)$ times, since d counts as four-fold on k sheets of $\Gamma_{4[n(k+1)-1]}$; and $\gamma_{4(n-1)}$ taken $k+1$ times, accounting for the intersection of $\gamma_{4(n-1)}$ and $\sigma_{n(k+2)-1}$, the conjugates of $\Gamma_{4[n(k+1)-1]}$ and $\sigma_{n(k+2)-1}$, respectively.

13. *Transforms of principal elements.* The transforms of the principal elements may be accounted for analytically. Since the transformation is involutorial, an $S_{2n(k+1)-1}$ of the web must go into a plane; d into $\tau_{2[n(k+1)-1]}$; $\gamma_{4(n-1)}$ into $\Gamma_{4[n(k+1)-1]}$; and since d is a consecutive double line on the given surface, the image of d, which is $\tau_{2[n(k+1)-1]}$, appears twice.

$$S_{2n(k+1)-1}: d^{2(n-2)(k+1)}k(n-2)\bar{d}^2\gamma_{4(n-1)}^{2k+1}[(6n-8)(k+1)-2]g.$$

The image of $\sigma_{n(k+2)-1}$ under I is itself; the image of $\tau_{2[n(k+1)-1]}$ is d; and the image of $\Gamma_{4[n(k+1)-1]}$ is $\gamma_{4(n-1)}$.

Thus, the intersections of the principal surfaces and the transforms of the principal elements confirm all the details of the Table of Characteristics.

CORNELL UNIVERSITY.

MINIMAL SURFACES OF UNIPLANAR DERIVATION.

By E. F. BECKENBACH.*

If the coördinates of a surface

$$x_j = x_j(u, v), \qquad (j = 1, 2, 3), \tag{1}$$

defined in the circle

$$R = [(u - u_0)^2 + (v - v_0)^2]^{1/2} < \rho, \tag{2}$$

are expressed in terms of isothermic parameters, that is, if

$$E = G, \quad F = 0, \tag{3}$$

then a necessary and sufficient condition that the surface be minimal is that the functions (1) be harmonic.

This being so, these functions are the real parts of analytic functions,

$$x_j = \Re f_j(z), \qquad (z = u + iv). \tag{4}$$

Now (3) becomes

$$\sum_{j=1}^{3} f'_j{}^2 = 0. \tag{5}$$

If the surface lies on the plane

$$x_3 = 0,$$

equation (5) becomes equivalent to the Cauchy-Riemann differential equations involving x_1 and x_2. Consequently, a map of (2) given by an analytic function of a complex variable is a special case of an isothermic map of (2) on a minimal surface. Since, further, minimal surfaces possess many of the properties of maps given by analytic functions, minimal surfaces well might be said to constitute the space analogue of maps given by analytic functions.

The purpose of this paper is to define, and to point out some of the properties of, a class of minimal surfaces which seems to correspond closely to that class of maps of (2), given by analytic functions of the complex variable z, which lie on a single-sheeted plane, or, as we shall say, on uniplanar † regions. An equivalent characterization of these functions is that they take on no value more than once in (2). For a discussion of such functions, see, among others, L. Bieberbach, *Lehrbuch der Funktionentheorie*,

* National Research Fellow.

† German *schlicht*.

Vol. II, pp. 82-94 (1927), or L. R. Ford, *Automorphic Functions*, pp. 169-177 (1929).

If and only if it is possible so to choose the coördinate axes that the functions

$$(6) \qquad M(z) = f_1 + if_2, \quad N(z) = f_1 - if_2$$

both map (2) *on uniplanar regions, we say that the functions* (1) *map* (2) *on a minimal surface of uniplanar derivation.*

This is a large but by no means exhaustive class of minimal surfaces. For example, the functions

$$x_1 = \Re\left[z + \tfrac{1}{4}z^2\right],$$

$$x_2 = \Re\left[\frac{i}{4}z^2\right],$$

$$x_3 = \Re\left[\frac{2i}{3}(1+z)^{3/2}\right],$$

map the unit circle $|z| < 1$ on a minimal surface of uniplanar derivation, so that, as we shall show, this surface possesses many of the properties of uniplanar analytic maps; but the surface does not lie on a plane.

If the functions (6) give uniplanar maps with one choice of coördinates, it does not follow that necessarily they give uniplanar maps with all choices. Thus, for the functions

$$f_1 = z, \quad f_2 = 0, \quad f_3 = -iz$$

we have

$$M(z) = z \text{ and } N(z) = z$$

both mapping (2) on uniplanar regions; but

$$f_1 - if_3 = 0.$$

We shall assume in the future that unless otherwise stated the coördinates of a minimal surface of uniplanar derivation have been so chosen that $M(z)$ and $N(z)$ both give uniplanar maps of the region of definition.

The following theorem will help justify our definition.

A uniplanar analytic map of (2) *is a minimal surface of uniplanar derivation. Conversely, if a minimal surface of uniplanar derivation lies on a Riemann surface, it lies on a uniplanar region.*

Suppose first that we have a uniplanar analytic map of (2). As we have pointed out, this is a special case of a minimal surface. If we choose coördinates so that the map lies in the (x_1, x_3)-plane, then

$$f_1 = x_1 + ix_3,$$
$$f_2 = 0,$$
$$f_3 = x_3 - ix_1 = -if_1,$$

and the functions (6) both give the uniplanar map in question. Consequently, the map is a minimal surface of uniplanar derivation.

Suppose secondly that we have (2) mapped conformally on a minimal surface lying on a Riemann surface but not on a single-sheeted plane. We shall show that the surface is not of uniplanar derivation. With the same axes as in the preceding case, the map is given by the analytic function

$$f_1 = x_1 + ix_3.$$

We must show that for any transformation of the rectangular axes,

$$X_j = \sum_{k=1}^{3} a_{j,k}x_k, \qquad (j = 1, 2, 3),$$

at least one of the new functions (6) does not give a uniplanar map.

The surface now is given by

$$X_j = a_{j,1}x_1(u, v) + a_{j,3}x_3(u, v) = \Re F_j(z).$$

Here

$$F_j = a_{j,1}f_1 + a_{j,3}f_3 = (a_{j,1} - ia_{j,3})f_1,$$

so that

$$F_1 + iF_2 = (a_{1,1} - ia_{1,3} + ia_{2,1} + a_{2,3})f_1,$$
$$F_1 - iF_2 = (a_{1,1} - ia_{1,3} - ia_{2,1} - a_{2,3})f_1.$$

Neither of these functions gives a uniplanar map, since f_1 does not give such a map. The surface, therefore, is not of uniplanar derivation.

It is an interesting but obvious observation that

a) *Every member of the family of real minimal surfaces associate to a minimal surface of uniplanar derivation is of uniplanar derivation.*

b) *Every real surface homothetic to a minimal surface of uniplanar derivation is a minimal surface of uniplanar derivation.*

If the original surface is given by (4), then the associate real minimal surfaces are given by

$$y_j = \Re e^{it}f_j(z),$$

where the parameter t is real. This family consists, to within their positions in space, of all the real minimal surfaces applicable to the first surface.

The homothetic real surfaces are given, to within their positions in space, by

$$y_j = \Re c f_j(z),$$

where the parameter c is real and not zero.

Combining the two families, we have that

$$(7) \qquad y_j = \Re c e^{it} f_j(z) = \Re w f_j(z),$$

where w is a non-vanishing complex parameter, define the two parameter family of minimal surfaces of uniplanar derivation applicable, or applicable after a magnification, to the minimal surface of uniplanar derivation defined by (4).

Obviously these surfaces (7) are minimal, for the y_j are the real parts of analytic functions and so are harmonic and, by (5),

$$w^2 \sum_{j=1}^{3} f'^2_j = 0.$$

Finally, the condition that they be of uniplanar derivation is that the functions

$$w(f_1 + if_2), \qquad w(f_1 - if_2)$$

give uniplanar maps, a condition that is satisfied since both of (6) give uniplanar maps, and since $w \neq 0$.

Every simply connected subregion of a minimal surface of uniplanar derivation is of uniplanar derivation.

Let the isothermic harmonic functions (1) map the circle (2) on the minimal surface of uniplanar derivation S, and consider any simply connected subregion S_1 of S. In this mapping, a subregion S'_1 of (2) is mapped on S_1. Let the function

$$z = z(Z), \qquad Z = U + iV,$$

map the circle

$$(2') \qquad [(U - U_0)^2 + (V - V_0)^2]^{1/2} < \rho$$

on S'_1. Since the functions (6) give uniplanar maps of (2), they give uniplanar maps of any part of (2), in particular of S'_1. Consequently, the functions

$$x_j = \Re f_j(z) = \Re f_j[z(Z)] = \Re F_j(Z),$$

which map (2′) on S_1, are such that

$$F_1 + iF_2 \quad \text{and} \quad F_1 - iF_2$$

give uniplanar maps of (2′).

If an analytic function, not identically constant, has a zero derivative

at any point, this point is a branch point of the function and the map is not uniplanar. Only at such points does the conformal character of the mapping break down. Similarly,

If the isothermic harmonic functions (1) *map the circle* (2) *on a minimal surface of uniplanar derivation, the mapping is conformal throughout.*

The mapping is conformal except where

$$E = 0,$$

that is, where simultaneously

$$f'_1 = 0, \quad f'_2 = 0, \quad f'_3 = 0,$$

since

$$E = \tfrac{1}{2} \sum_{j=1}^{3} | f'_j |^2.$$

The points of (2) where this condition is satisfied necessarily are isolated for any minimal surface. But if there are such points, the functions (6) do not give uniplanar maps and consequently the surface is not of uniplanar derivation.

The existence of such a singular point is a sufficient, but not a necessary, condition that a map be not uniplanar or of uniplanar derivation.

The equations of Weierstrass for a general minimal surface,

$$x_1 = \Re \int i\,[g^2(z) - h^2(z)]\, dz,$$

$$x_2 = \Re \int [g^2(z) + h^2(z)]\, dz,$$

$$x_3 = \Re \int 2ig(z)h(z)\, dz,$$

are obtained from (5) by means of the equations of definition

$$f'_1 + if'_2 = 2ig^2(z),$$
$$f'_1 - if'_2 = -2ih^2(z).$$

Hence, for minimal surfaces of uniplanar derivation,

$$\phi(z) = \int g^2(z)\, dz = (1/2i) M(z), \tag{8}$$
$$\psi(z) = \int h^2(z)\, dz = (-1/2i) N(z),$$

both give uniplanar maps.

A direct computation gives

$$E = [g(z)\bar{g}(z) + h(z)\bar{h}(z)]^2,$$

where $\bar{\theta}$ denotes the conjugate imaginary of θ.

The length of a curve on the minimal surface is given by

$$L = \int_{z_0}^{z_1} E^{1/2} \mid dz \mid = \int_{z_0}^{z_1} \mid g \mid^2 \mid dz \mid + \int_{z_0}^{z_1} \mid h \mid^2 \mid dz \mid .$$

The lengths of curves on the maps

$$w_1 = \phi(z) \quad \text{and} \quad w_2 = \psi(z)$$

are given by

$$L_1 = \int_{z_0}^{z_1} \mid g \mid^2 \mid dz \mid \quad \text{and} \quad L_2 = \int_{z_0}^{z_1} \mid h \mid^2 \mid dz \mid$$

respectively. Consequently, if the paths in (2) are the same in the three cases,

$$L = L_1 + L_2.$$

The following theorem is analogous to Bieberbach's famous theorem on uniplanar maps.*

If the isothermic harmonic functions (1) *map the circle* (2) *on a finite surface of uniplanar derivation, then the minimum distance on the surface from the image of* (u_0, v_0) *to the boundary is greater than or at least equal to* $E_0{}^{1/2}\rho/4$, *where* E_0 *designates the value of* E *at* (u_0, v_0). *No closer inequality holds for all surfaces of this type.*

The above distance is given by the minimum, for all paths of integration, of the integral

$$\int_{R=0}^{\rho} E^{1/2} \mid dz \mid = \int_{R=0}^{\rho} \mid g \mid^2 \mid dz \mid + \int_{R=0}^{\rho} \mid h \mid^2 \mid dz \mid .$$

Since the functions $\phi(z)$ and $\psi(z)$ give finite uniplanar maps of (2), we can apply Bieberbach's theorem to these functions, getting, for any path of integration,

$$(9) \qquad \begin{aligned} \int_{R=0}^{\rho} \mid g \mid^2 \mid dz \mid &\geq \mid g_0 \mid^2 \rho/4, \\ \int_{R=0}^{\rho} \mid h \mid^2 \mid dz \mid &\geq \mid h_0 \mid^2 \rho/4, \end{aligned}$$

where g_0 and h_0 are the values of g and h at (u_0, v_0).

* Bieberbach, *loc. cit.*, p. 86; Ford, *loc. cit.*, p. 169.

Adding the two inequalities (9) and applying

$$E_0^{1/2} = |g_0|^2 + |h_0|^2,$$

we obtain the desired inequality

$$\int_{R=0}^{\rho} E^{1/2} |dz| \geq E_0^{1/2} \rho/4. \tag{10}$$

Since, by Bieberbach's theorem, the inequalities (9) are the closest possible for all finite uniplanar analytic maps, and since uniplanar maps are a special case of maps of uniplanar derivation, it follows that (10) is the closest inequality possible for all finite minimal surfaces of uniplanar derivation.

The equality in the preceding theorem,

$$\textit{minimum} \int_{R=0}^{\rho} E^{1/2} |dz| = E_0^{1/2} \rho/4, \tag{11}$$

can hold only if the map of uniplanar derivation is actually uniplanar.

For (11) to hold, each equality in (9) must hold for a certain path of integration; furthermore, this path must be the same in both cases. Consequently, $\phi(z)$ and $\psi(z)$ are derived from Koebe's function

$$W = Z/(1 + e^{-i\alpha}Z)^2, \qquad z = z_0 + \rho Z,$$

by means of

$$W = [\phi(z_0 + \rho Z) - \phi(z_0)]/\rho g_0^2,$$

and

$$W = [\psi(z_0 + \rho Z) - \psi(z_0)]/\rho h_0^2,$$

respectively, where the real parameter α has the same value in both cases and where

$$z_0 = u_0 + iv_0.$$

It is no restriction on the generality to take

$$\phi(z_0) = 0, \qquad \psi(z_0) = 0,$$

so that

$$\psi(z) = (h_0^2/g_0^2)\phi(z). \tag{12}$$

From (6), (8), and (12),

$$f_1 + if_2 = 2i\phi,$$
$$f_1 - if_2 = -2i(h_0^2/g_0^2)\phi,$$

so that, by (5),

$$
(13) \qquad \begin{aligned}
f_1 &= i(1 - h_0^2/g_0^2)\phi, \\
f_2 &= (1 + h_0^2/g_0^2)\phi, \\
f_3 &= \pm 2i(h_0/g_0)\phi.
\end{aligned}
$$

Setting

$$
\begin{aligned}
\phi &= A(u, v) + iB(u, v) = A + iB, \\
h_0/g_0 &= a + ib,
\end{aligned}
$$

and using (4), we find from (13) that

$$
\begin{aligned}
x_1 &= 2abA - (1 - a^2 + b^2)B, \\
x_2 &= (1 + a^2 - b^2)A - 2abB, \\
x_3 &= \mp 2bA \mp 2aB.
\end{aligned}
$$

The transformation of the rectangular coördinates defined by

$$
\begin{aligned}
X_1 &= \frac{b}{(a^2 + b^2)^{1/2}} x_1 + \frac{a}{(a^2 + b^2)^{1/2}} x_2, \\
X_2 &= \frac{-a(1 - a^2 - b^2)}{(a^2 + b^2)^{1/2}(1 + a^2 + b^2)} x_1 + \frac{b(1 - a^2 - b^2)}{(a^2 + b^2)^{1/2}(1 + a^2 + b^2)} x_2 \\
&\qquad \mp \frac{2(a^2 + b^2)}{(a^2 + b^2)^{1/2}(1 + a^2 + b^2)} x_3, \\
X_3 &= \frac{2a}{1 + a^2 + b^2} x_1 - \frac{2b}{1 + a^2 + b^2} x_2 \pm \frac{a^2 + b^2 - 1}{1 + a^2 + b^2} x_3,
\end{aligned}
$$

preserves orthogonality of axes and yields

$$
\begin{aligned}
X_1(u, v) &= (aA - bB) \frac{1 + a^2 + b^2}{(a^2 + b^2)^{1/2}}, \\
X_2(u, v) &= (bA + aB) \frac{1 + a^2 + b^2}{(a^2 + b^2)^{1/2}}, \\
X_3(u, v) &= 0;
\end{aligned}
$$

whence,

$$
(14) \qquad \begin{aligned}
X_1 + iX_2 &= \frac{1 + a^2 + b^2}{(a^2 + b^2)^{1/2}} (a + ib)[A(u, v) + iB(u, v)] \\
&= E_0^{1/2} e^{i(\theta_1 + \theta_2)} \rho W,
\end{aligned}
$$

where

$$
g_0 = e^{i\theta_1} | g_0 |, \qquad h_0 = e^{i\theta_2} | h_0 | .
$$

The function (14) gives a uniplanar analytic map of (2), since W gives such a map.

We now establish the following deformation theorem.

If the isothermic harmonic functions (1) *map the circle* (2) *on a finite*

surface of uniplanar derivation, then at any point within the circle the following inequalities hold,

$$(15) \qquad \frac{1-r}{(1+r)^3} \leq \left[\frac{E(u,v)}{E_0}\right]^{1/2} \leq \frac{1+r}{(1-r)^3},$$

where

$$(16) \qquad [(u-u_0)^2+(v-v_0)^2]^{1/2} = r\rho.$$

Further, no closer inequalities hold for all surfaces of this type.

We know that *

$$(17) \qquad \begin{gathered} \frac{1-r}{(1+r)^3} \leq \frac{|g(u,v)|^2}{|g_0|^2} \leq \frac{1+r}{(1-r)^3}, \\ \frac{1-r}{(1+r)^3} \leq \frac{|h(u,v)|^2}{|h_0|^2} \leq \frac{1+r}{(1-r)^3}. \end{gathered}$$

It is a simple algebraic fact that

$$\frac{|g|^2+|h|^2}{|g_0|^2+|h_0|^2}$$

lies between

$$|g|^2/|g_0|^2 \quad \text{and} \quad |h|^2/|h_0|^2,$$

whence (15).

The equalities in (17) and (15) hold only for those functions for which the equalities hold in the preceding theorem. The conclusions follow as before that no closer inequalities hold for all surfaces of this type and that if the equalities hold the map is uniplanar, given by (14).

If (u_1, v_1) and (u_2, v_2) lie at the same distance $r\rho$ from (u_0, v_0), we obtain from (15) the inequality

$$\left(\frac{1-r}{1+r}\right)^4 \leq \left[\frac{E(u_1,v_1)}{E(u_2,v_2)}\right]^{1/2} \leq \left(\frac{1+r}{1-r}\right)^4.$$

If the isothermic harmonic functions (1) *map the circle* (2) *on a finite surface of uniplanar derivation, then the minimum distance $m(u,v)$ on the surface from the image of (u_0, v_0) to the point corresponding to (u,v) satisfies the inequalities*

$$(18) \qquad [r/(1+r)^2]\,\rho E_0^{1/2} \leq m(u,v) \leq [r/(1-r)^2]\,\rho E_0^{1/2},$$

where r is given by (16). *No closer inequalities hold for all surfaces of this type.*

---

* Bieberbach, *loc. cit.*, p. 88; Ford, *loc. cit.*, pp. 172-173.

The second inequality of (15) gives, for any path of integration,

$$\int_{R=0}^{r} E^{1/2} \mid dz \mid \leq \int_{R=0}^{r} E_0^{1/2} [(1+r)/(1-r)^3] \mid dz \mid .$$

The first of these integrals, taken along a radius, is at least as great as $m(u, v)$. Hence,

$$(19) \quad m(u,v) \leq \int_0^r [(1+r)/(1-r)^3] \, \rho E_0^{1/2} \, dr = [r/(1-r)^2] \, \rho E_0^{1/2}.$$

If the curve C minimizes the distance, then

$$m(u,v) = \int_C E^{1/2} \mid dz \mid .$$

This integral is not increased if we substitute for

$$\mid dz \mid = [\rho^2 dr^2 + \rho^2 r^2 d\theta^2]^{1/2}$$

the not larger quantity $\rho \mid dr \mid$:

$$m(u,v) \geq \int_C \rho E^{1/2} \mid dr \mid .$$

The first inequality of (15) gives, then,

$$(20) \qquad m(u,v) \geq \int_C [(1-r)/(1+r)^3] \, \rho E_0^{1/2} \mid dr \mid$$

$$\geq \int_0^r [(1-r)/(1+r)^3] \, \rho E_0^{1/2} \, dr$$

$$= [r/(1+r)^2] \, \rho E_0^{1/2}.$$

From (19) and (20) we get (18). The equalities hold only as in the preceding two theorems.

These last two theorems yield the following pair of theorems. The proofs, depending on a division of the (u, v) region into squares of sufficiently small size, are exactly the same as those of the corresponding theorems concerning uniplanar maps, and therefore are not given here.*

Let Σ' be a uniplanar finite region and let Σ be a closed point set consisting only of interior points of Σ'. Let the isothermic harmonic functions (1) *map Σ' on a finite surface of uniplanar derivation. Then there exists a constant K, dependent on Σ and Σ' but independent of the mapping functions, such that if (u_1, v_1) and (u_2, v_2) are any two points of Σ, then*

* Bieberbach, *loc. cit.*, p. 89; Ford, *loc. cit.*, pp. 175-177.

$$\frac{1}{K} < \frac{E(u_1, v_1)}{E(u_2, v_2)} < K.$$

Finally,

In the mapping of the preceding theorem there exists a constant L, independent of the mapping functions, such that if

$$(u_1, v_1), \quad (u_2, v_2), \quad \text{and} \quad (u_3, v_3)$$

are any three points of Σ, then

$$m\,[(u_1, v_1),\ (u_2, v_2)] < L\,[E(u_3, v_3)]^{1/2},$$

where

$$m\,[(u_1, v_1),\ (u_2, v_2)]$$

denotes the minimum distance on the surface between the points (u_1, v_1) and (u_2, v_2).

THE RICE INSTITUTE,
PRINCETON UNIVERSITY.

CONCERNING REGULAR PSEUDO *D*-CYCLIC SETS.

By LEONARD M. BLUMENTHAL.

1. *Introduction.* A set of points is called pseudo d-cyclic provided that each three of the points is congruent with three points of a circle of metric diameter d, while the whole set is not congruent to a subset of the circle.* In a recent paper † the writer has completely characterized those pseudo d-cyclic sets that (1) contain no convex tripod, and (2) contain no pseudo d-cyclic quadruples that are pseudo-linear.‡ Such pseudo d-cyclic sets were called *proper*. The purpose of this paper is to characterize pseudo d-cyclic sets that contain no convex tripods, the assumption that the set contains no pseudo-linear quadruples that are pseudo d-cyclic not being made. We call such sets *regular*. The principal theorem of this paper proves that these sets are equilateral provided that they contain more than four points.

It has been shown in the paper referred to above that pseudo d-cyclic quadruples are of three kinds; namely, (1) pseudo d-cyclic quadruples that contain no linear triples, (2) pseudo d-cyclic quadruples that contain exactly three linear triples, and (3) pseudo d-cyclic quadruples that have all four of the triples they contain linear. The second case can occur only when the quadruple forms a convex tripod, and hence this case is excluded from this discussion. The third case occurs when the quadruple is pseudo-linear, and no two of its points are diametral. We shall refer to pseudo d-cyclic quadruples as being of the *first, second,* or *third kinds,* depending upon whether they are in the first, second, or third of the above classifications, respectively. Further, if p_1, p_2, p_3, p_4 is a pseudo d-cyclic quadruple of either the first or third kind, then $p_1p_2 = p_3p_4, p_1p_3 = p_2p_4, p_2p_3 = p_1p_4$.

* Throughout this paper the word "circle" refers to the circumference. Distance between two points of a circle is defined to be the length of the shorter arc of the circle joining them.

† "A Complete Characterization of Proper Pseudo D-Cyclic Sets of Points," *American Journal of Mathematics*, Vol. 54 (1932), pp. 387-396.

‡ Four points form a convex tripod provided that one of the points lies *between* each of the three pairs of points contained in the remaining three points. (A point q is said to lie between two points p, r provided that $pq + qr = pr$. The triple p, q, r is said to be *linear*. We shall denote the above relation by the notation pqr.) Four points form a pseudo-linear set provided that each three of the points is congruent to three points of a line, while the four points are not congruent to four points of a line.

2. In this section, we establish two lemmas and a theorem concerning pseudo d-cyclic quadruples and quintuples.

LEMMA 1. *A pseudo d-cyclic quadruple does not contain two diametral points.*

To prove this, we suppose that two points, say p_1, p_3, of the pseudo d-cyclic quadruple p_1, p_2, p_3, p_4 are diametral; that is, the distance p_1p_3 equals d. Then evidently the two triples p_1, p_2, p_3 and p_1, p_4, p_3 containing these two points are linear, and we have $p_1p_2 + p_2p_3 = p_1p_3$; $p_1p_4 + p_4p_3 = p_1p_3$. Now the quadruple is either of the second or the third kind (since it contains linear triples). The supposition that the quadruple is of the third kind leads immediately to a contradiction, for a quadruple with four linear triples (which are d-cyclic) and two points with a distance equal to d may be imbedded in a circle of metric diameter d, and hence is d-cyclic. Suppose, then, that the quadruple is of the second kind. Then the four points must form a convex tripod; *i. e.*, exactly three of the triples contained in the four points are linear, and one of the points lies between three pairs of points. But from the above two relations, p_2 lies between p_1 and p_3, while p_4 lies between p_1 and p_3. Hence, the four points do not form a convex tripod, and this case is also impossible. Hence, the lemma is proved.

LEMMA 2. *If a pseudo d-cyclic quintuple contains two diametral points, and a pseudo-linear quadruple which is pseudo d-cyclic, then the quintuple contains a convex tripod.*

Let p_1, p_2, p_3, p_4, p_5 be a quintuple satisfying the conditions of the lemma. We may choose the labeling so that p_2, p_3, p_4, p_5 is the pseudo-linear quadruple which is pseudo d-cyclic. Then the pair of diametral points that, by hypothesis, the quintuple contains, is not contained in this quadruple, since, by Lemma 1, a pseudo d-cyclic quadruple does not contain a pair of diametral points. We may letter the points so that p_1, p_5 are diametral.

Since the quadruple p_2, p_3, p_4, p_5 is pseudo-linear, we have $p_2p_3 = p_4p_5 = a$; $p_3p_4 = p_2p_5 = b$; $p_2p_4 = p_3p_5 = c$; and all four of the triples contained in the quadruple are linear. Then,

$$(a + b - c)(a - b + c)(-a + b + c) = 0.$$

Now, $p_1p_5 = d$, and hence we have $p_1p_2 + p_2p_5 = d$; $p_1p_4 + p_4p_5 = d$; $p_1p_3 + p_3p_5 = d$. Combining these relations with the ones above, we obtain $p_1p_2 = d - b$; $p_1p_3 = d - c$; $p_1p_4 = d - a$.

Consider, now the quadruple p_1, p_2, p_3, p_4. We have three cases to consider.

Case A. $a + b = c$. The point p_3 of the quadruple p_1, p_2, p_3, p_4 is to lie between the pairs p_1, p_2; p_1, p_4; p_2, p_4, while the triple p_1, p_2, p_4 i linear. Thus, the quadruple forms a convex tripod.

Case B. $a + c = b$. The point p_2 lies between each of the three of points p_1, p_3; p_1, p_4; p_3, p_4, and the quadruple p_1, p_2, p_3, p_4 is again se form a convex tripod.

Case C. $b + c = a$. Here, the point p_4 lies between three pairs of p and the quadruple is a convex tripod.

Thus, in any one of the three cases, the quintuple contains a cc tripod, and the lemma is proved.

These two lemmas enable us to prove the following theorem, of v much use is to be made:

THEOREM I. *If a pseudo d-cyclic quintuple contains a pair of diam points, then the quintuple contains a convex tripod.*

Assume the labeling so that, as before, the points p_1, p_5 are diam We have, then, $p_1p_2 + p_2p_5 = d$; $p_1p_4 + p_4p_5 = d$; $p_1p_3 + p_3p_5 = d$. since the circle has the congruence order four, at least one of the quadr contained in the five points is pseudo d-cyclic. By Lemma 1, this quad does not contain the points p_1, p_5. We may choose the labeling so tha quadruple p_2, p_3, p_4, p_5 is pseudo d-cyclic. If this quadruple is pseudo-li the theorem is proved by an application of Lemma 2. Suppose, then, tha quadruple is not pseudo-linear. If it is a quadruple of the second kind theorem is proved. We suppose, finally, that the quadruple is of the kind. Then we have the relations $p_2p_3 = p_4p_5 = a$; $p_3p_4 = p_2p_5 = b$; $= p_3p_5 = c$; and $a + b + c = 2d$. Consider the quadruple p_1, p_2, p_3, p is seen at once that the point p_1 lies between each of the three pairs of p p_2, p_3; p_2, p_4; p_3, p_4, and the quadruple forms a convex tripod. Hence theorem is proved.

We give an example of a pseudo d-cyclic quintuple containing a ps linear quadruple that is pseudo d-cyclic, a pair of diametral points, a convex tripod. The ten distances determined by the five points are give means of the table:

	p_1	p_2	p_3	p_4	p_5
p_1	0	$d/3$	$d/2$	$d/6$	d
p_2	$d/3$	0	$d/6$	$d/2$	$2d/3$
p_3	$d/2$	$d/6$	0	$d/3$	$d/2$
p_4	$d/6$	$d/2$	$d/3$	0	$5d/6$
p_5	d	$2d/3$	$d/2$	$5d/6$	0

This quintuple contains three d-cyclic quadruples. The pseudo-linear quadruple that is also pseudo d-cyclic is p_1, p_2, p_3, p_4, and the convex tripod is formed by the four points p_2, p_3, p_4, p_5.

3. We now establish certain lemmas concerning regular pseudo d-cyclic quintuples (that is, pseudo d-cyclic quintuples no four points of which form a convex tripod) that will enable us to prove the principal theorem of this paper characterizing regular pseudo d-cyclic sets containing more than four points.

LEMMA 1. *A regular pseudo d-cyclic quintuple does not contain exactly one d-cyclic quadruple.*

We suppose that the regular pseudo d-cyclic quintuple p_1, p_2, p_3, p_4, p_5 contains just a single d-cyclic quadruple. We may choose the labeling so that this quadruple is p_1, p_2, p_3, p_4. Then the remaining four quadruples are pseudo d-cyclic of the first or third kinds (since the quintuple is regular). In either case, if we let p_i, p_j, p_k, p_n represent any one of these quadruples, we have $p_ip_j = p_kp_n$; $p_ip_k = p_jp_n$; $p_jp_k = p_ip_n$. Writing these relations out for each of the four quadruples, it is seen that all of the ten distances determined by the five points are equal. Then, in particular, the quadruple p_1, p_2, p_3, p_4 is equilateral, and hence is not d-cyclic, contrary to the hypothesis. Thus, the lemma is proved.

LEMMA 2. *A regular pseudo d-cyclic quintuple does not contain exactly two d-cyclic quadruples.*

We suppose that the regular pseudo d-cyclic quintuple p_1, p_2, p_3, p_4, p_5 contains two (exactly) d-cyclic quadruples, and we select the labeling of the points so that these two d-cyclic quadruples are p_1, p_2, p_3, p_4 and p_2, p_3, p_4, p_5. Then the remaining three quadruples are pseudo d-cyclic of the first or third kinds, and, in either case, we have the following relations:

$$\text{(A)} \quad p_1p_2 = p_1p_3 = p_1p_4 = p_2p_5 = p_3p_5 = p_4p_5; \quad p_2p_3 = p_2p_4 = p_3p_4 = p_1p_5.$$

Since the quadruple p_1, p_2, p_3, p_4 is d-cyclic, it contains at least two linear triples. Now the triple p_2, p_3, p_4 is equilateral, and hence, it is not linear. Then two of the remaining three triples, p_1, p_2, p_3; p_1, p_2, p_4; p_1, p_3, p_4 are linear. We have, thus, three cases to consider. All three cases are treated similarly to the one discussed here.

Case A. The triples p_1, p_2, p_3; p_1, p_2, p_4 are linear. Then relations (A) yield that $p_2p_1p_3$ and $p_2p_1p_4$ exist. But then, we have $p_3p_1p_4$, and the four points form a convex tripod, and hence are not d-cyclic, contrary to the hypothesis.

Thus, the lemma is proved.

LEMMA 3. *A regular pseudo d-cyclic quintuple does not contain exactly three d-cyclic quadruples.*

We suppose the contrary, and select the labeling so that p_1, p_2, p_3, p_4; p_2, p_3, p_4, p_5 are the two pseudo d-cyclic quadruples. Then we have the relations $p_1p_2 = p_3p_4 = p_2p_5$; $p_1p_3 = p_2p_4 = p_3p_5$; $p_2p_3 = p_1p_4 = p_4p_5$. We distinguish two cases:

Case A. The pseudo d-cyclic quadruples p_1, p_2, p_3, p_4; p_2, p_3, p_4, p_5 are both of the first kind. The three quadruples p_1, p_2, p_3, p_5; p_1, p_2, p_4, p_5; p_1, p_3, p_4, p_5 are, by hypothesis, d-cyclic, and hence each contains at least two linear triples. Since the two quadruples p_1, p_2, p_3, p_4; p_2, p_3, p_4, p_5 are supposed pseudo d-cyclic of the first kind, none of the seven distinct triples that they contain are linear. Thus, of the ten triples contained in the five points, only three of them are linear. These triples are p_1, p_2, p_5; p_1, p_3, p_5; p_1, p_4, p_5. Each of the three d-cyclic quadruples must contain two of these triples. From the above relations, we obtain that $p_1p_2p_5$; $p_1p_3p_5$; $p_1p_4p_5$ exist; *i. e.*, $p_1p_5 = 2(p_1p_2) = 2(p_1p_3) = 2(p_1p_4)$. Then $p_1p_2 = p_2p_3 = p_1p_3$, and since the triple p_1, p_2, p_3 is d-cyclic, each of these distances equals $2d/3$. But then $p_1p_5 = 4d/3$, which is impossible, for since any triple containing p_1, p_5 is d-cyclic, the distance p_1p_5 cannot exceed d. Hence this case is not possible.

Case B. One of the quadruples p_1, p_2, p_3, p_4; p_2, p_3, p_4, p_5 is of the third kind. Then both of the quadruples are of the third kind; for, since the quintuple is regular, pseudo d-cyclic quadruples are either of the first or third kinds. The triple p_2, p_3, p_4 is common to both of the quadruples, and the supposition that one quadruple is of the third kind means that this triple is linear. Then the other quadruple contains a linear triple, and hence it is of the third kind. Then the seven distinct triples contained in these two quadruples are linear, and in addition, we have, as before, the relations $p_1p_2 = p_3p_4 = p_2p_5$; $p_1p_3 = p_2p_4 = p_3p_5$; $p_2p_3 = p_1p_4 = p_4p_5$. Now, we may label the three points common to the two quadruples so that $p_2p_3p_4$ exists. Applying the above relations, we have $p_1p_2p_3$; $p_3p_4p_1$; $p_4p_1p_2$; $p_3p_4p_5$; $p_4p_5p_2$; $p_3p_2p_5$ holding. Consider the d-cyclic quadruple p_1, p_2, p_3, p_5. Two of its triples are linear in the order $p_1p_2p_3$ and $p_3p_2p_5$. Three sub-cases present themselves.

Sub-case 1. The remaining two triples contained in p_1, p_2, p_3, p_5 are not linear. Since these two triples are d-cyclic we have $p_1p_3 + p_3p_5 + p_5p_1 = 2d$, $p_1p_2 + p_2p_5 + p_5p_1 = 2d$. Subtracting the second relation from the first,

and making use of the fact that $p_3p_2p_5$ exists, we obtain $p_1p_3 = p_1p_2 - p_2p_3$, which is impossible, since $p_1p_2p_3$ exists. Hence, this sub-case cannot occur.

Sub-case 2. Only one of the two triples p_1, p_3, p_5; p_1, p_2, p_5 is linear. Suppose that p_1, p_3, p_5 is linear. Then $p_1p_3p_5$ exists, and since p_1, p_2, p_5 is not linear, we have $p_1p_2 + p_2p_5 + p_5p_1 = 2d$. From these relations, it is easily seen that $p_1p_5 = d + p_2p_3$, which is impossible. Hence this cannot occur.

Suppose that p_1, p_2, p_5 is selected as the only one of the two triples to be linear. Then we obtain $p_1p_2p_5$, and the point p_2 is seen to lie between p_1p_5; p_1p_3; p_3p_5; *i. e.*, the points p_1, p_2, p_3, p_5 form a convex tripod and hence are not d-cyclic, as supposed.

Sub-case 3. Both of the triples p_1, p_2, p_5; p_1, p_3, p_5 are linear. Then $p_1p_2p_3$; $p_3p_2p_5$; $p_1p_3p_5$; $p_1p_2p_5$ exist. From these relations, it is immediate that $p_1p_2 = p_1p_3$, which is impossible, for $p_1p_2p_3$ exists. A contradiction in the form $p_1p_2 = p_1p_3 + p_3p_2$ may also be found.

Hence, none of the three sub-cases under Case B can occur, and the theorem is proved.

LEMMA 4. *A regular pseudo d-cyclic quintuple does not contain exactly one pseudo d-cyclic quadruple.*

The proof of this final lemma is divided into two parts.

Part 1. We suppose that the regular pseudo d-cyclic quintuple, p_1, p_2, p_3, p_4, p_5 contains exactly one pseudo d-cyclic quadruple, and we assume that this quadruple is of the first kind. We select the labeling so that this quadruple is p_1, p_2, p_3, p_4. Then none of the four triples contained in this quadruple is linear. The quadruples p_1, p_2, p_3, p_5; p_1, p_2, p_4, p_5; p_1, p_3, p_4, p_5; p_2, p_3, p_4, p_5 are each d-cyclic and hence each contains at least two linear triples. Since each of these quadruples contains a non-linear triple from the quadruple p_1, p_2, p_3, p_4, they can contain at most three linear triples. But if a d-cyclic quadruple contains three linear triples, while the fourth triple is not linear, the quadruple must contain two diametral points.* But then, according to the theorem of section 2, the quintuple would contain a convex tripod, which is impossible, since the quintuple is, by hypothesis, regular. Hence, each of the four d-cyclic quadruples contains exactly two linear triples. We may select the triples p_1, p_3, p_5; p_2, p_3, p_5 in the quadruple p_1, p_2, p_3, p_5 to be linear. Then, in order that one of the remaining three quadruples may not

* L. M. Blumenthal, "A Complete Characterization of Proper Pseudo D-Cyclic Sets of Points," *American Journal of Mathematics*, Vol. 54 (1932), p. 393.

have three linear triples, we must select p_1, p_4, p_5; p_2, p_4, p_5 as linear. Then each of the four d-cyclic quadruples contains exactly two linear triples. The triples p_1, p_2, p_5; p_3, p_4, p_5 are not linear, but since they are d-cyclic, we have the relations $p_1p_2 + p_2p_5 + p_5p_1 = 2d$; $p_3p_4 + p_4p_5 + p_5p_3 = 2d$. Also, from the fact that the quadruple p_1, p_2, p_3, p_4 is pseudo d-cyclic, we have $p_1p_2 = p_3p_4$; $p_1p_4 = p_2p_3$; $p_1p_3 = p_2p_4$. From the linearity of the four triples, we have the relations:

$$(\mathrm{A})\quad \begin{aligned}
&(p_1p_3 + p_3p_5 - p_5p_1)(p_1p_3 - p_3p_5 + p_5p_1)(-p_1p_3 + p_3p_5 + p_5p_1) = 0\\
&(p_2p_3 + p_3p_5 - p_5p_2)(p_2p_3 - p_3p_5 + p_5p_2)(-p_2p_3 + p_3p_5 + p_5p_2) = 0\\
&(p_1p_4 + p_4p_5 - p_5p_1)(p_1p_4 - p_4p_5 + p_5p_1)(-p_1p_4 + p_4p_5 + p_5p_1) = 0\\
&(p_2p_4 + p_4p_5 - p_5p_2)(p_2p_4 - p_4p_5 + p_5p_2)(-p_2p_4 + p_4p_5 + p_5p_2) = 0
\end{aligned}$$

An examination of the first two of the relations (A) together with the relations that immediately precede them, leads to the result that either $p_1p_3p_5$ and $p_3p_5p_2$ or $p_1p_5p_3$ and $p_2p_3p_5$ exist. A similar examination of the last two of the relations (A) yields the fact that either $p_1p_4p_5$ and $p_4p_5p_2$ or $p_1p_5p_4$ and $p_2p_4p_5$ exist. But it is easily shown that no one of the four cases obtained by grouping these two pairs of relations is possible. Hence the case considered by Part 1 of the theorem cannot occur.*

Part 2. In this part we assume that the regular pseudo d-cyclic quintuple p_1, p_2, p_3, p_4, p_5 contains exactly one pseudo d-cyclic quadruple, and this quadruple is pseudo-linear. We may label the points so that the quadruple p_1, p_2, p_3, p_4 is pseudo-linear, and so that $p_1p_2p_3$, $p_2p_3p_4$, $p_3p_4p_1$, $p_4p_1p_2$ exist.

Consider, now, any one of the four d-cyclic quadruples, say p_1, p_2, p_4, p_5

---

* To see, for example, that of the nine combinations that may be obtained from the first two of the relations (A) only the two combinations given above are consistent with the relations preceding (A), it is sufficient to examine each of these nine combinations. To indicate the method employed to reject combinations, we consider the case of $p_1p_3p_5$, $p_2p_3p_5$; i. e.,

(1) $$p_1p_3 + p_3p_5 = p_1p_5,$$

(2) $$p_2p_3 + p_3p_5 = p_2p_5.$$

From these two relations we obtain

(3) $$p_1p_5 - p_2p_5 = p_1p_3 - p_2p_3.$$

Now the triples p_1, p_2, p_3; p_1, p_2, p_5 are d-cyclic, *not* linear. Whence, we have

(4) $$p_1p_2 + p_2p_3 + p_3p_1 = 2d,$$

(5) $$p_1p_2 + p_2p_5 + p_5p_1 = 2d.$$

From (3) and (4) we obtain, $2(p_1p_3) = 2d - p_1p_2 - p_2p_5 + p_1p_5$, and using (5), we get $2(p_1p_3) = 2(p_1p_5)$. But this, together with (1) implies that $p_3p_5 = 0$, which is impossible, since the points p_3 and p_5 are distinct.

contained in the quintuple. This quadruple being d-cyclic, and not containing two diametral points (by reason of Lemma 2, section 2), must contain either exactly two or exactly four linear triples. Two cases present themselves.

Case A. The quadruple p_1, p_2, p_4, p_5 contains four linear triples. Consider the d-cyclic quadruple p_2, p_3, p_4, p_5. This quadruple contains the two linear triples p_2, p_3, p_4 and p_2, p_4, p_5. But, since $p_2p_3p_4$ exists, the linearity of p_2, p_4, p_5 in any order is possible only if all four of the triples contained in this quadruple are linear. Then evidently all ten of the triples contained in the quintuple are linear. In a similar manner, this result is seen to follow if any of the four d-cyclic quadruples is assumed to have four linear triples.

Case B. Each of the four d-cyclic quadruples has exactly two linear triples. The quadruples p_1, p_2, p_4, p_5; p_1, p_3, p_4, p_5; p_1, p_2, p_3, p_5; p_2, p_3, p_4, p_5 contain the linear triples p_1, p_2, p_4; p_1, p_3, p_4; p_1, p_2, p_3; p_2, p_3, p_4, respectively, and each contains exactly one other linear triple. It is found that this is possible only if one of the three following combinations are selected as linear:

(1) p_1, p_2, p_5; p_3, p_4, p_5 (2) p_1, p_3, p_5; p_2, p_4, p_5 (3) p_2, p_3, p_5; p_1, p_4, p_5.

We examine the first combination. The quadruple p_1, p_2, p_3, p_5 contains only the two linear triples p_1, p_2, p_3 and p_1, p_2, p_5. Since $p_1p_2p_3$ exists, the second triple is linear in the order $p_2p_1p_5$. Consider, now, the quadruple p_1, p_2, p_4, p_5. It contains the two linear triples p_1, p_2, p_4; p_1, p_2, p_5, and since $p_2p_1p_4$ and $p_2p_1p_5$ exist, all of the triples contained in this quadruple are linear, contrary to the hypothesis for Case B. Hence combination (1) cannot occur.

To show the impossibility of combination (2), consider the quadruple p_1, p_2, p_3, p_5. Since $p_1p_2p_3$ exists, the linearity of p_1, p_3, p_5 in any order demands the linearity of all four of its triples.

Finally, we consider the combination (3). From the quadruple p_1, p_3, p_4, p_5 we obtain $p_3p_4p_1$ and $p_4p_1p_5$. Then the quadruple p_1, p_2, p_4, p_5 is seen to have all of its triples linear, since $p_2p_1p_4$ and $p_4p_1p_5$ exist. Thus, the assumption that each of the quadruples contained in the quintuple has exactly two linear triples leads to a contradiction. Hence, Case A alone is possible, and all ten triples contained in the five points are linear.

But the straight line is known to have the quasi-congruence order 3 *; *i. e.*, any set of points containing more than four points such that each triple is congruent to three points of a line, is congruent to a sub-set of the line.

* Karl Menger, "New Foundation of Euclidean Geometry," *American Journal of Mathematics*, Vol. 53 (1931), p. 727.

The five points p_1, p_2, p_3, p_4, p_5 have been shown to have all of their triples linear, and hence they are congruent with five points of a line. This, however, is impossible, for the five points contain the quadruple p_1, p_2, p_3, p_4, which, by hypothesis is pseudo-linear. Hence the five points cannot be congruent with five points of a line.

Thus, Case A having been shown to be impossible, the theorem is proved.

Since the circle has the congruence order 4, all of the quadruples contained in a pseudo d-cyclic quintuple cannot be d-cyclic. This fact, together with the four lemmas proved above, establishes the following theorem:

THEOREM II. *A regular pseudo d-cyclic quintuple does not contain any d-cyclic quadruples.*

Hence, if a set of five points is such that (1) all ten of its triples are d-cyclic, (2) the set does not contain a convex tripod, (3) one of its quadruples is d-cyclic, then the set is congruent with five points of a circle of metric diameter d.

We have immediately the theorem characterizing regular pseudo d-cyclic quintuples.

THEOREM III. *A regular pseudo d-cyclic quintuple is equilateral.*

Since the quintuple does not contain any d-cyclic quadruples, all of its quadruples are pseudo d-cyclic of either the first or third kinds. In either case, the quadruple has its "opposite" distances equal. Writing these relations for each of the five quadruples, we have immediately that all ten of the distances determined by the five points are equal. Hence the quintuple is equilateral. Since each triple is d-cyclic, we note that each of the ten distances equals $2d/3$.

COROLLARY 1. *All quadruples of a regular pseudo d-cyclic quintuple are pseudo d-cyclic of the first kind.*

COROLLARY 2. *A regular pseudo d-cyclic quintuple does not contain any linear triples.*

4. We consider now regular pseudo d-cyclic sets containing more than four points. For these sets we prove the following theorem.

THEOREM IV. *A regular pseudo d-cyclic set containing more than four points is equilateral.*

We have seen that this theorem is true if the set consists of exactly five points. To prove the theorem, we shall assume it true for a set consisting of

k points $(k > 4)$, and show that it follows that it is true for sets containing $k+1$ points.

Let $p_1, p_2, \cdots, p_k, p_{k+1}$ be a regular pseudo d-cyclic set containing exactly $k+1$ points. This set contains at least one regular pseudo d-cyclic set of k points; for, otherwise, every set of k points contained in the $k+1$ points would be d-cyclic, and since $k > 4$, every quadruple contained in the $k+1$ points would be d-cyclic, *a fortiori*. But, since the circle has the congruence order 4, the $k+1$ points would be d-cyclic, instead of pseudo d-cyclic, as supposed.

We may assume the labeling so that $p_1, p_2, \cdots, p_k$ is pseudo d-cyclic. We now show that at least one other set of k points contained in the $k+1$ points is pseudo d-cyclic. Suppose that each of the other k sets of k points contained in the $k+1$ points is d-cyclic. Then all of the quadruples contained in these k sets are d-cyclic. But all of the quadruples to be found in the pseudo d-cyclic set of k points, $p_1, p_2, \cdots, p_k$, are contained in the remaining k sets, and since these quadruples are all d-cyclic, the k points $p_1, p_2, \cdots, p_k$ have all their quadruples d-cyclic and hence are themselves d-cyclic, which contradicts the previous assumption. Hence the $k+1$ points contain another pseudo d-cyclic set of k points. We may select the labeling so that this set is $p_2, p_3, \cdots, p_k, p_{k+1}$. The two pseudo d-cyclic sets of k points shown to be contained in the $k+1$ points are regular, and since the theorem is assumed true for regular pseudo d-cyclic sets of k points, we have that each of these two sets is equilateral, with all of their distances equal to $2d/3$. Thus, of the $(1/2)k(k+1)$ distances determined by the $k+1$ points, all are seen to be equal except the distance p_1p_{k+1}, which does not enter into the above two sets. To determine this distance, consider any triple, say p_1, p_2, p_{k+1}, containing this pair of points. This triple is not linear, for if it were, then

$$(p_1p_2 + p_2p_{k+1} - p_1p_{k+1})(p_1p_2 - p_2p_{k+1} + p_1p_{k+1})(-p_1p_2 + p_2p_{k+1} + p_1p_{k+1}) = 0$$

and since $p_1p_2 = p_2p_{k+1} = 2d/3$, we have $p_1p_{k+1} = 4d/3$, which is impossible. Hence the triple p_1, p_2, p_{k+1} is not linear, and since it is d-cyclic, we have $p_1p_2 + p_2p_{k+1} + p_1p_{k+1} = 2d$, from which $p_1p_{k+1} = 2d/3$. Thus, the set of $k+1$ points is equilateral, and the theorem is proved.

Two corollaries similar to those following the theorem characterizing regular pseudo d-cyclic quintuples may be stated.

THE RICE INSTITUTE.

ON ARRAYS OF NUMBERS.*

By LEONARD CARLITZ.†

1. *Introduction.* This paper is concerned with the investigation of arrays of numbers, $a_{n,s}$, defined by a relation of the type

$$(1) \qquad a_{n+1,s} = \sum_{i=0}^{N} \beta_i(s,n)\, a_{n,s-N+i},$$

—where N is independent of both s and n,—plus a set of "initial" conditions such as

$$a_{1,1} = 1; \qquad a_{1,s} = 0 \quad \text{for} \quad s \neq 1.‡$$

Thus, for example, for $N = 1$, $\beta_0 \equiv \beta_1 \equiv 1$, we have the Pascal Triangle—that is, the array of combinatorial coefficients; for $N = 1$, $\beta_0 \equiv 1$, $\beta_1 = s$, the array of Stirling numbers.

The numbers $a_{n,s}$ are best studied by means of the operator of order N,

$$(2) \qquad \beta_0 E^N + \beta_1 E^{N-1} + \cdots + \beta_N,$$

where E is the well known symbol defined by

$$Ef(s) = f(s-1);$$

we suppose in the following that E operates only on s. We shall be interested in determining "explicit" expressions for the elements of the array (1). As will appear in § 6 fairly simple expressions may be obtained when the operator (2) satisfies certain conditions. Before deriving actual expressions for the constituents of an arbitrary array, we first set up certain arrays that seem to possess some interest in themselves (§§ 2, 3, 4).§

In § 7 some of the results deduced in the earlier parts of the paper are applied to the evaluation of certain finite sums. The consideration of one of these sums indicates a connection between particular arrays and Bernoulli polynomials in several variables.

2. *Expansion of* $(x^{\lambda+\mu}D^{\mu})^n$. It is evident that we may put (μ integer $\geqq 0$)

---

* Presented to the American Mathematical Society, November 29, 1930.

† National Research Fellow.

‡ Precisely these conditions hold in all cases studied in the present paper.

§ A number of special cases are treated by the writer in the *American Mathematical Monthly*, Vol. 37 (1930), pp. 472-479.

$$(3) \qquad (x^{\lambda+\mu}D^{\mu})^{n} = \sum_{s=1}^{\mu(n-1)+1} a_{n,s}(x^{s+\mu-1+n\lambda}D^{s+\mu-1}),$$

where of course $D \equiv d/dx$. In order to determine the $a_{n,s}$ we apply $(x^{\lambda+\mu}D^{\mu})$ to both sides of (3). Then

$$(x^{\lambda+\mu}D^{\mu})^{n+1} = \sum a_{n,s}(x^{\lambda+\mu}D^{\mu})(x^{s+\mu-1+n\lambda}D^{s+\mu-1})$$
$$= \sum_{s} a_{n,s} \sum_{j=0}^{\mu} \binom{\mu}{j} (s+\mu-1+n\lambda)\,!^{\,j}\, x^{s+2\mu-j-1+(n+1)\lambda}D^{s+2\mu-j-1},$$

where

$$\binom{\mu}{j} = \frac{\mu\,!}{j\,!(\mu-j)\,!}; \qquad \mu\,!^{\,j} = \mu(\mu-1)\cdots(\mu-j+1), \quad \mu\,!^{\,0} = 1.$$

Writing t for $s+\mu-j$,

$$(x^{\lambda+\mu}D^{\mu})^{n+1} = \sum_{t=1}^{\mu n+1} \sum_{j=0}^{\mu} a_{n,t-\mu+j} \binom{\mu}{j} (t+j-1+n\lambda)\,!^{\,j}(x^{t+\mu-1+(n+1)\lambda}D^{t+\mu-1}).$$

Comparison with (3) shows that

$$(4) \qquad a_{n+1,s} = \sum_{j=0}^{\mu} \binom{\mu}{j} (s+j-1+n\lambda)\,!^{\,j}\, a_{n,s-\mu+j}$$
$$= \sum \binom{\mu}{j} (s+j-1+n\lambda)\,!^{\,j}\, E^{\mu-j}\cdot a_{n,s}.$$

Evidently the $a_{n,s}$ satisfy the initial conditions of § 1 and hence define an array of numbers of the type to be considered. The associated operator is

$$\sum_{j} \binom{\mu}{j} (s+j-1+n\lambda)\,!^{\,j}\, E^{\mu-j};$$

it is easily verified that this can be written

$$(5) \quad = (E+s+n\lambda)(E+s+1+n\lambda)\cdots(E+s+\mu-1+n\lambda),$$

and that the operators in (5) are all commutative.

3. *Expansion of* $(x^{\mu}D^{\lambda+\mu})^{n}$. We suppose now that λ and μ are both non-negative integers. By the results of the preceding section the expansion of the operator $(x^{\mu}D^{\lambda+\mu})^{n}$ would appear to require an array of *order* (i. e. the order of the associated operator) $\lambda+\mu$; however it is not difficult to show that an array of order μ will suffice. Accordingly we put

$$(6) \qquad (x^{\mu}D^{\lambda+\mu})^{n} = \sum_{s=1}^{\mu(n-1)+1} b_{n,s}(x^{s+\mu-1}D^{s+\mu-1+n\lambda}).$$

As above apply $x^{\mu}D^{\lambda+\mu}$ to both sides of (6):

$$(x^{\mu}D^{\lambda+\mu})^{n+1} = \sum^{s} b_{n,s} \sum_{j=0}^{\lambda+\mu} \binom{\lambda+\mu}{j} (s+\mu-1)\,!^{\,j}\,(x^{s+2\mu-j-1}D^{s+2\mu-j-1+(n+1)\lambda})$$
$$= \sum_{t}\sum_{j} \binom{\lambda+\mu}{j} (t+j-1)\,!^{\,j}\, b_{n,t-\mu+j}(x^{t+\mu-1}D^{t+\mu-1+(n+1)\lambda}).$$

Comparing with (6),

$$(7) \qquad b_{n+1,s} = \sum_{j=0}^{\lambda+\mu} \binom{\lambda+\mu}{j} (s+j-1)\,!^{\,j}\, b_{n,s-\mu+j}$$
$$= \sum \binom{\lambda+\mu}{j} (s+j-1)\,!^{\,j}\, E^{\mu-j} \cdot b_{n,s}.$$

However this is not of the form (1); and indeed it involves an operator of order $\lambda+\mu$. In order to attain the desired result we notice first that the operator in (7)

$$= (E+s)(E+s+1)\cdots(E+s+\lambda+\mu-1)E^{-\lambda},$$

where as in (5) the factors are commutative. Then

$$b_{n+1,s} = (E+s)\cdots(E+s+\lambda+\mu-1)E^{-\lambda} b_{n,s}$$
$$= \prod_{h=0}^{\lambda+\mu-1}(E+s+h)\cdot E^{-\lambda}\prod_{k=0}^{\lambda+\mu-1}(E+s+k)\cdot E^{-\lambda} b_{n-1,s}$$
$$= \prod (E+s+h)\cdot \prod (E+s+\lambda+k)\cdot E^{-2\lambda} b_{n-1,s}$$
$$(8) \qquad = \prod_{h=0}^{\lambda+\mu-1}\prod_{k=0}^{n-1}(E+s+h+k\lambda)\cdot E^{-n\lambda} b_{1,s}.$$

Now

$$\prod_{j=1}^{n\lambda}(E+s+j-1)\cdot E^{-n\lambda} = \sum_{j=0}^{n\lambda}\binom{n\lambda}{j}(s+j-1)\,!^{\,j}\, E^{-j};$$

and putting

$$\prod_{h=0}^{\mu-1}\prod_{k=1}^{n}(E+s+h+k\lambda) = \sum_{k=0}^{n\mu} A_h E^h$$

(8) becomes

$$\sum_{h=0}^{n\mu}\sum_{j=0}^{n\lambda}\binom{n\lambda}{j} A_h (s+j-h-1)\,!^{\,j}\, b_{1,s+j-h}.$$

But $b_{1,s+j-h}$ vanishes except when $s+j-h=1$, and then $(s+j-h-1)\,!^{\,j}$ vanishes unless $j=0$. Therefore

$$b_{n+1,s} = \prod_{h=0}^{\mu-1}\prod_{k=1}^{n}(E+s+h+k\lambda)\cdot b_{1,s},$$

and finally

$$(9) \qquad b_{n+1,s} = \prod_{h=1}^{\mu}(E+s+h-1+n\lambda)\cdot b_{n,s},$$

which defines an array of order μ; indeed it is identical with that defined by (4) of § 2.

4. *Arrays associated with*

$$m\,!^{\,\lambda+\mu}(m-\lambda)\,!^{\,\lambda+\mu}\cdots(m-[n-1]\lambda)\,!^{\,\lambda+\mu}.$$

We begin with the proof of the identity

$$(10)\quad m\,!^{\lambda+\mu}(m-n\lambda)\,!^{\lambda+\mu} = \sum_{j=0}^{\lambda+\mu}\Lambda_j(k,n)\,(m+k+j-1)\,!^{j}(m-[n+1]\lambda-\mu-k+1)\,!^{j}$$

where

$$\Lambda_j(k)=\Lambda_j(k,n) = \binom{\lambda+\mu}{j}(k+\lambda+\mu-1)\,!^{\lambda+\mu-j}(k+[n+1]\lambda+\mu-1)\,!^{\lambda+\mu-j}$$

for all integral $k \geqq 1-\lambda-\mu$.

Evidently (10) holds for $k=1-\lambda-\mu$; assuming then that it holds for all larger values up to and including k, we show without difficulty that it holds for $k+1$. Since

$$(m+k)(m-[n+1]\lambda-\mu-k+1) = (m+k+j)(m-[n+1]\lambda-\mu-k-j-1) + j([n+1]\lambda+\mu+2k+j-1),$$

the right side of (10)

$$= \sum_{j=0}^{\lambda+\mu}\Lambda_j(k)\Big[(m+k+j)\,!^{j}(m-[n+1]\lambda-\mu-k)\,!^{j} + j([n+1]\lambda+\mu+2k+j-1)(m+k+j-1)\,!^{j-1}(m-[n+1]\lambda-\mu-k)\,!$$

$$= \sum_{j=0}^{\lambda+\mu}\{\Lambda_j(k)+(j+1)([n+1]\lambda+\mu+2k+j)\Lambda_{j+1}(k)\} \times (m+k+j)\,!^{j}(m-[n+1]\lambda-\mu-k)\,!^{j}.$$

But the quantity within the { }

$$= \binom{\lambda+\mu}{j}(k+\lambda+\mu-1)\,!^{\lambda+\mu-j}(k+[n+1]\lambda+\mu-1)\,!^{\lambda+\mu-j} + (j+1)([n+1]\lambda+\mu+2k+j)\binom{\lambda+\mu}{j+1}(k+\lambda+\mu-1)\,!^{\lambda+\mu-j-1} \times (k+[n+1]\lambda+\mu-1)\,!^{\lambda+\mu-j-1}$$

$$= \binom{\lambda+\mu}{j}(k+\lambda+\mu-1)\,!^{\lambda+\mu-j-1}(k+[n+1]\lambda+\mu-1)\,!^{\lambda+\mu-j-1} \cdot \big[(k+j)(k+j+n\lambda)+(\lambda+\mu-j)([n+1]\lambda+\mu+2k+j)\big]$$

$$= \binom{\lambda+\mu}{j}(k+\lambda+\mu)\,!^{\lambda+\mu-j}(k+[n+1]\lambda+\mu)\,!^{\lambda+\mu-j}$$

completing the induction.

4.1. We now consider the expansion

$$(11)\ldots\quad m\,!^{\lambda+\mu}(m-\lambda)\,!^{\lambda+\mu}\cdots(m-2(n-1)\lambda)\,!^{\lambda+\mu} = \sum_{s=1}^{\mu n+1} c_{n,s}(m+s-1)\,!^{(2n-1)\lambda+2s+\mu-2};$$

that such an expansion exists may be shown by repeated application of the obvious identity

$$m(m-h)=(m+k)(m-h-k)+(h+k)k.$$

To determine the $c_{n,s}$, put $m-\lambda$ in place of m in (11); and then multiply the left side of (11) by

(12) $$m\,!^{\lambda+\mu}(m-2n\lambda)\,!^{\lambda+\mu}$$

and the right side by the expression obtained by applying (10) to (12):

$$\begin{aligned} m\,!^{\lambda+\mu}\cdots(m-2n\lambda)\,!^{\lambda+\mu} &= \sum_{s=1}^{\mu(n-1)+1} c_{n,s}(m+s-\lambda-1)\,!^{(2n-1)\lambda+2s+\mu-2} \\ &\quad\cdot\sum_{j=0}^{\lambda+\mu}\Lambda_j(s-\lambda)\cdot(m+s-\lambda+j-1)\,!^{j}(m-2n\lambda-\mu-s+1)\,!^{j} \\ &=\sum_s\sum_j \Lambda_j(s-\lambda,2n)\,c_{n,s}(m+s-\lambda+j-1)\,!^{(2n-1)\lambda+2s+2j+\mu-2} \\ &=\sum_{t=1}^{\mu n+1}\ \sum_{j=0}^{\lambda+\mu}\Lambda_j(t-j,2n)\,c_{n,t-j+\lambda}(m+t-1)\,!^{(2n+1)\lambda+2t+\mu-2}. \end{aligned}$$

Comparison with (11) shows that

(13) $$\begin{aligned} c_{n+1,s} &= \sum \Lambda_j(s-j,2n)\,c_{n,s-j+\lambda} \\ &= \sum_{j=0}^{\lambda+\mu}\Lambda_j(s-j,2n)E^{j-\lambda}\cdot c_{n,s}. \end{aligned}$$

We now proceed exactly as at the corresponding point in § 3; we may prove by an easy induction

$$\sum_{j=0}^{\lambda+\mu}\Lambda_j(s-j,2n)E^j=\prod_{j=1}^{\lambda+\mu}\Big[E+(s+j-1)(s+[2n+1]\lambda+\mu-j)\Big],$$

the factors on the right being permutable.* (13) then becomes

$$\begin{aligned} c_{n+1,s} &= \prod_{j=1}^{\lambda+\mu}\Big[E+(s+j-1)(s+[2n+1]\lambda+\mu-j)\Big]E^{-\lambda}\cdot c_{n,s} \\ &= \prod_{j=1}^{\lambda+\mu}\Big[E+(s+j-1)(s+[2n+1]\lambda+\mu-j)\Big] \\ &\quad\cdot\prod_{i=1}^{\lambda+\mu}[E+(s+\lambda+i-1)(s+2n\lambda+\mu-i)]\cdot E^{-2\lambda}\cdot c_{n-1,s} \end{aligned}$$

* It will be noticed that in the following factorization of the same operator the factors are *not* permutable:

$$(E+s(s+2n\lambda))\,(E+(s+1)(s+2n\lambda+1))$$
$$\cdots(E+(s+\lambda+\mu-1)(s+[2n+1]\lambda+\mu-1)).$$

$$(14)\quad = \prod_{j=1}^{\lambda+\mu} \prod_{h=0}^{n-1} \Big[E + (s + h\lambda + j - 1)(s + [2n - h + 1]\lambda + \mu - j) \Big] \cdot E^{-n\lambda} \cdot c_{1,s}.$$

But

$$\prod_{j=1}^{n\mu} \Big[E + (s + j - 1)(s + [2n+1]\lambda + \mu - j) \Big] \cdot E^{-n\lambda}$$
$$= \sum_{j=0}^{n\mu} \binom{n\lambda}{j} (s + j - 1)\,!^{j} (s + [n+1]\lambda + \mu + j - 1)\,!^{j} E^{-j},$$

and, if we put

$$\prod_{j=1}^{\mu} \prod_{h=1}^{n} \Big[E + (s + h\lambda + j - 1)(s + [2n - h + 1]\lambda + \mu - j) \Big] = \sum_{h=0}^{\mu n} A_h E^h,$$

(14) becomes

$$\sum_{h,j} A_h \binom{n\lambda}{j} (s - h + j - 1)\,!^{j} (s + [n+1]\lambda + \mu + j - h - 1)\,!^{j} E^{h-j} \cdot c_{1,s},$$

so that all terms vanish except those for which $s - h + j = 1$; and the presence of $(s - h + j - 1)\,!^{j}$ necessitates $j = 0$. Finally, therefore, (14) may be rewritten as

$$(15)\quad c_{n+1,s} = \prod_{j=1}^{\mu} \prod_{h=1}^{n} \Big[E + (s + h + j - 1)(s + [2n - h + 1]\lambda + \mu - j) \Big] \cdot c_{1,}$$

while this does not reduce to (1), yet as will appear from the results of the next section, from this form an explicit expression for the $c_{n,s}$ is easily obtained (see (36)).

4. 2. If we assume μ to be even, it is possible to replace (11) by an expression holding when the number of factorials in the left member of (11) is either odd or even. This expansion is (writing 2ν for μ)

$$(16)\quad m\,!^{\lambda+2\nu} \cdots (m - [n-1]\lambda)\,!^{\lambda+2\nu} = \sum_{s=1}^{\nu(n-1)+1} d_{n,s} (m + s - 1)\,!^{2(s+\nu-1)+n\lambda}.$$

For n odd this reduces to (11); assume then $n = 2p$. We proceed exactly as above. In (16) put $m - \lambda$ in place of m, multiply the two sides of the resulting equality by the corresponding members of (applying (10))

$$m\,!^{\lambda+2\nu} (m - [2p+1]\lambda)\,!^{\lambda+2\nu}$$
$$= \sum_{j=0}^{\lambda+2\nu} \Delta_j(k, 2p+1)(m + k + j - 1)\,!^{j} (m - [2p+2]\lambda - 2\nu - k + 1)\,!^{j}.$$

Then without difficulty it is seen that

$$d_{2p+2,s} = \sum_{j=0}^{\lambda+2\nu} \binom{\lambda+2\nu}{j} (s+j-1)\,!^{\,j}(s+j-1+[2p+1]\lambda)\,!^{\,j}\, d_{2p,s-2\nu+j}$$

$$= \prod_{j=1}^{\lambda+2\nu}\Big[E+(s+j-1)(s+[2p+2]\lambda+2\nu-j)\Big]\cdot E^{-\lambda}\cdot d_{2p,s}$$

$$= \prod_{j=1}^{\lambda+2\nu}\Big[E+(s+j-1)(s+[2p+2]\lambda+2\nu-j)\Big]$$

$$\times\Big[E+(s+\lambda+j-1)(s+[2p+1]\lambda+2\nu-j)\Big]\cdot E^{-2\lambda}d_{2p-2,s}$$

$$(17)\qquad = \prod_{j=1}^{\lambda+2\nu}\prod_{h=0}^{p-1}\Big[E+(s+h\lambda+j-1) \times (s+[2p-h+2]\lambda+2\nu-j)\Big]\cdot E^{-p\lambda}\, d_{2,s}.$$

As for $d_{2,s}$, starting with the identity like (10),

$$(m-\lambda)\,!^{\,2\nu} = \sum_{i=0}^{\nu}\binom{\nu}{i}(\lambda+2\nu)\,!^{\,\nu-i}(\lambda+\nu)\,!^{\,\nu-i}(m+i)\,!^{\,i}(m-2\lambda-2\nu)\,!^{\,i},$$

we see that

$$m\,!^{\,\lambda+2\nu}(m-\lambda)\,!^{\,\lambda+2\nu} = \sum_{i=0}^{\nu}\binom{\nu}{i}(\lambda+2\nu)\,!^{\,\nu-i}(\lambda+\nu)\,!^{\,\nu-i}(m+i)\,!^{\,2i+2\lambda+2\nu},$$

whence

$$d_{2,j+1} = \binom{\nu}{j}(\lambda+2\nu)\,!^{\,\nu-j}(\lambda+\nu)\,!^{\,\nu-j};$$

or what amounts to the same thing,

$$d_{2,s} = \prod_{j=1}^{\nu}[E+(s+\lambda+j-1)(s+\lambda+2\nu-j)]\; d_{1,s}.$$

Substituting this into (17),

$$d_{2p+2,s} = \prod_{j=1}^{\lambda+2\nu}\prod_{h=0}^{p-1}\Big[E+(s+h\lambda+j-1)(s+[2p-h+2]\lambda+2\nu-j)\Big]$$

$$(18)\quad \cdot\prod_{i=1}^{\nu}\Big[E+(s+[p+1]\lambda+i-1)(s+[p+1]\lambda+2\nu-i)\Big]\cdot E^{-p\lambda}d_{1,s}.$$

Noting that the brackets in the right of (18) are permutable, we may evidently apply the method of reduction already used several times. In this way we get

$$(19)\quad d_{2p,s} = \prod_{j=1}^{2\nu}\prod_{h=1}^{p-1}\Big[E+(s+h\lambda+j-1)(s+[2p-h]\lambda+2\nu-j)\Big]$$

$$\cdot\prod_{j=1}^{\nu}\Big[E+(s+p\lambda+j-1)(s+p\lambda+2\nu-j)\Big]\cdot d_{1,s}.$$

Now split the double product into two parts,

$$\prod_{j=1}^{2\nu}\prod_{h=1}^{p-1}=\prod_{j=1}^{\nu}\prod_{h}\cdot\prod_{j=\nu+1}^{2\nu}\prod_{h}=\Pi_1\cdot\Pi_2,$$

say; in the double product Π_2 replace j by $2\nu-j+1$, h by $2p-h$. Then

$$\Pi_1\,\Pi_2=\prod_{j=1}^{\nu}\prod_{\substack{h=1\\ h\neq p}}^{2p-1}\Big[E+(s+h\lambda+j-1)(s+[2p-h]\lambda+2\nu-j)\Big],$$

and finally (19) becomes

$$(20)\quad d_{2p,s}=\prod_{j=1}^{\nu}\prod_{h=1}^{2p-1}\Big[E+(s+h\lambda+j-1)(s+[2p-h]\lambda+2\nu-j)\Big]\cdot d_{1,s}.$$

If n in (16) $=2p-1$, comparison with (11) shows that $d_{2p-1,s}=c_{p,s}$, which by (15)

$$=\prod_{j=1}^{2\nu}\prod_{h=1}^{p-1}\Big[E+(s+h\lambda+j-1)(s+[2p-h-1]\lambda+2\nu-j)\Big]\cdot c_{1,s},$$

and this, exactly as (19) was transformed into (20), becomes

$$(21)\quad d_{2p-1,s}=\prod_{j=1}^{\nu}\prod_{h=1}^{2p-2}\Big[E+(s+h\lambda+j-1)(s+[2p-h-1]+2\nu-j)\Big]\cdot d_{1,}$$

Finally (20) and (21) may be written

$$(22)\quad d_{n,s}=\prod_{j=1}^{\nu}\prod_{h=1}^{n-1}\Big[E+(s+h\lambda+j-1)(s+[n-h]\lambda+2\nu-j)\Big]\cdot d_{1,s},$$

which is the formula sought.

5. *Explicit expressions.* If $\Omega(n)$ denote the operator

$$A_0(s,n)E^N+\cdots+A_N(s,n),$$

then evidently, from

$$a_{n+1,s}=\Omega(n)a_{n,s},$$

we get

$$(23)\qquad a_{n+1,s}=\Omega(n)\Omega(n-1)\cdots\Omega(1)a_{1,s};$$

hence if the product of operators on the right be expanded into a polynomial in E, the coefficients thus obtained will furnish the value of $a_{n+1,s}$. We shall limit ourselves in the following to the case in which $\Omega(n)$ can be split into a product of permutable linear factors. It will be noticed that all the special cases treated above actually lead to operators $\Omega(n)$ of this type.

We first investigate the conditions under which an operator

$$\Omega=E^k+\cdots+A_k(s)$$

may have this property. If we define $B_\nu(s)$ by the equation

$$(24)\qquad A_\nu(s) - A_\nu(s-1) = A_{\nu-1}(s)B_{\nu-1}(s) \qquad (\nu = 1,\cdots,k),$$

(where $A_0(s) \equiv 1$) then we shall prove that *a necessary and sufficient condition that* Ω *be a product of permutable linear operators is furnished by*

$$(25)\qquad B_{k-\nu}(s) = B_{k-1}(s) + B_{k-1}(s-1) + \cdots + B_{k-1}(s-\nu+1) \qquad (\nu = 2, 3, \cdots, k).$$

To prove the *necessity*, we remark first that two permutable linear operators must be of the form

$$E + \alpha(s), \qquad E + \alpha(s) + \mu,$$

where μ is free of s. If then

$$\Omega = \prod_{\nu=1}^{k} (E + \alpha(s) + \mu_\nu),$$

it is clear that

$$(E + \alpha(s))\Omega = \Omega(E + \alpha(s)),$$

and therefore

$$\Omega E - E\Omega = \alpha(s)\Omega - \Omega\alpha(s).$$

But

$$(26)\qquad \Omega E - E\Omega = \sum (A_{k-\nu}(s) - A_{k-\nu}(s-1))E^{\nu+1},$$

and

$$\alpha(s)\Omega - \Omega\alpha(s) = \sum A_{k-\nu}(s)[\alpha(s) - \alpha(s-\nu)]\,E^\nu.$$

Comparing coefficients we find that

$$A_{k-\nu+1}(s) - A_{k-\nu+1}(s-1) = A_{k-\nu}(s)[\alpha(s) - \alpha(s-\nu)];$$

therefore, by (24),

$$B_\nu(s) = \alpha(s) - \alpha(s-k+\nu),$$

and (25) follows immediately.

To prove the *sufficiency*, write

$$\beta(s) = \sum_{\nu=1}^{s} B_{k-1}(\nu),$$

so that, using (25),

$$B_{k-\nu}(s) = \beta(s) - \beta(s-\nu) \qquad (\nu = 1,\cdots,k).$$

But writing equation (26) in the form

$$\Omega E - E\Omega = \sum A_{k-\nu}(s)B_{k-\nu}(s)E^\nu,$$

it is immediately apparent that

$$(27)\qquad (E + \beta(s))\Omega = \Omega(E + \beta(s)).$$

In order to complete the proof, we have then to prove the

LEMMA. *If an operator* Ω *satisfy equation* (27), *then it may be written as a product of permutable linear factors.*

If we put

$$\Omega_1 = \Omega - (E + \beta(s))^k,$$

it is evident that

$$(E + \beta(s))\Omega_1 = \Omega_1(E + \beta(s)),$$

so that the coefficient of E^{k-1} in Ω_1 is free of s; call it p_1. We next put

$$\Omega_2 = \Omega_1 - p_1(E + \beta(s))^{k-1},$$

and it appears in exactly the same way that p_2, the coefficient of E^{k-2} in Ω_2, is free of s. Continuing in this way we see that we may write

$$\Omega = (E + \beta(s))^k + p_1(E + \beta(s))^{k-1} + \cdots + p_k, \tag{28}$$

where the p_ν are all free of s. Therefore, finally if $\rho_1, \cdots, \rho_k$ are the roots of the equation

$$\rho^k + p_1\rho^{k-1} + \cdots + p_k = 0,$$

we have

$$\Omega = (E + \beta(s) + \rho_1) \cdots (E + \beta(s) + \rho_k),$$

completing the proof.

Returning to equation (23), we consider the expansion into a polynomial of

$$\Omega = \prod_{\nu=1}^{k} (E + \alpha_\nu(s, n)) = \sum_{i=0}^{k} A_i^{(k)}(s, n) E^i, \tag{29}$$

where

$$\alpha_\nu(s, n) = \alpha(s) + \rho_\nu(n). \tag{30}$$

We take first the case in which

$$\rho_\nu(n) \equiv 0 \qquad (\nu = 1, \cdots, k).$$

It is then fairly clear that $A_i^{(k)}$ may be written

$$A_i^{(k)}(s, n) = A_i^k(s) = \sum_{j=0}^{i} B_{ij}(s)\alpha^k(s - j). \tag{31}$$

To solve for B_{ij}, note that

$$A_i^i = 1, \quad A_i^k = 0 \quad \text{for} \quad k < i.$$

Then by Cramer's Rule, (31) yields

$$B_{ij}(s) = \frac{(-1)^j}{\prod_{\mu=0}^{j-1} [\alpha(s - \mu) - \alpha(s - j)] \prod_{\nu=j+1}^{i} [\alpha(s - j) - \alpha(s - \nu)]}, \tag{32}$$

after some easy transformations.

To treat the general case of $A_i^{(k)}(s,n)$ in equation (29), we note that, exactly as in (28), for the Ω now under discussion,

$$\Omega \equiv (E+\beta(s))^k + p_1(n)(E+\beta(s))^{k-1} + \cdots + p_k(n),$$

where $p_\nu(n)$ is the ν-th elementary symmetric function of the quantities $\rho_1(n), \cdots, \rho_k(n)$. Therefore, by (29) and (31), we find finally that

$$A_i^{(k)}(s,n) = \sum_{j=0}^{i} B_{ij}(s)\alpha_1(s-j,n)\cdots\alpha_k(s-j,n). \tag{33}$$

Equation (33), together with (32), furnishes the desired explicit expression for the coefficients in (29).

Hence, for the array

$$a^1_{n+1,s} = [E+\alpha(s)+\beta(n)]\cdot a^1_{n,s},$$

from

$$a^1_{n+1,s} = [E+\alpha(s)+\beta(n)]\cdots[E+\alpha(s)+\beta(1)]\cdot a^1_{1,s},$$

we derive immediately

$$a^1_{n,s} = \sum_{j=0}^{s-1} \frac{(-1)^j[\alpha(s-j)+\beta(1)]\cdots[\alpha(s-j)+\beta(n)]}{\prod_{h=0}^{j-1}[\alpha(s-h)-\alpha(s-j)]\prod_{t=j+1}^{s-1}[\alpha(s-j)-\alpha(s-t)]}.$$

For the array

$$a^\mu_{n+1,s} = [E+\alpha(s)+\beta_1(n)]\cdots[E+\alpha(s)+\beta_\mu(n)]\cdot a^\mu_{n,s}$$

we consider the operator

$$\prod_{i=1}^{\mu}\prod_{j=1}^{n}[E+\alpha(s)+\beta_i(j)].$$

Accordingly in (30) take

$$k=n\mu; \quad \rho_1(n)=\beta_1(1), \cdots, \rho_k(n)=\beta_\mu(n).$$

Then

$$a^\mu_{n+1,s} = \sum_{j=0}^{s-1} \frac{(-1)^j\prod\prod[\alpha(s-j)+\beta_h(t)]}{\prod[\alpha(s-h)-\alpha(s-j)]\prod[\alpha(s-j)-\alpha(s-t)]}$$

$$= \sum_{j=1}^{s} \frac{(-1)^{s-j}\prod_{h=1}^{\mu}\prod_{t=1}^{n}[\alpha(j)+\beta_h(t)]}{\prod_{h=j+1}^{s}[\alpha(h)-\alpha(j)]\prod_{t=1}^{j-1}[\alpha(j)-\alpha(t)]}, \tag{34}$$

which furnishes the desired explicit expression.

6. *Explicit expressions for the arrays of* §§ 1, 3. By means of (32) or

(34) we can immediately write down simple expressions for the elements of the arrays defined in §§ 1, 3. For (5), (9),

$$a_{n+1,s}=\prod_{h=1}^{\mu}[E+s+n\lambda+h-1]\cdot a_{n,s};$$

take

$$\alpha(s)=s,\qquad \beta_h(t)=t\lambda+h-1;$$

then

$$(35)\quad a_{n+1,s}=\frac{1}{(s-1)!}$$
$$\times\sum_{j=1}^{s}(-1)^{s-j}\binom{s-1}{j-1}(j+\lambda+\mu-1)!^{\mu}\cdots(j+n\lambda+\mu-1)!^{\mu}$$

the right side may be transformed into a sum depending on λ and μ, but it is unnecessary to consider that here.

Turning now to (15), we take in (30), (32), (33)

$$\alpha(s)=s^2+([2n-1]\lambda+\mu-1)s;$$
$$\rho_{(j-1)\mu+h}(n)=(j\lambda+h-1)([2n-j-1]\lambda+\mu-h)$$
$$(j=1,\cdots,n-1;\quad h=1,\cdots,\mu);$$

then by (34) after some easy reduction,

$$(36)\quad c_{n,s}=\frac{1}{(2s+[2n-1]\lambda+\mu-1)!}\sum_{j=1}^{s}(-1)^{s-j}\binom{2s+[2n-1]\lambda+\mu-1}{s-j}$$
$$\cdot(2j+[2n-1]\lambda+\mu-1)(j+[2n-1]\lambda-1)!^{[2n-1]\lambda}\prod_{h=1}^{2n-1}(j+h\lambda+\mu-1)!^{\mu}.$$

To determine the $d_{n,s}$ of (16) we make use of (22), and take

$$\alpha(s)=s^2+(n\lambda+2\nu-1)s;$$
$$\rho_{(j-1)\nu+h}(n)=(j\lambda+h-1)([n-j]\lambda+2\nu-h)$$
$$(j=1,\cdots,n-1;\quad h=1,\cdots,\nu).$$

Then exactly as in deriving (32),

$$(37)\quad d_{n,s}=\frac{1}{(2s+n\lambda+2\nu-1)!}\sum_{j=1}^{s}(-1)^{s-j}\binom{2s+n\lambda+2\nu-1}{s-j}$$
$$\cdot(2j+n\lambda+2\nu-1)(j+n\lambda-1)!^{n\lambda}\prod_{h=1}^{n}(j+h\lambda+2\nu-1)!^{2\nu}.$$

7. *Some applications.* If we operate on x^m with both members of (6) we find that

$$m\,!^{\lambda+\mu}\cdots(m-[n-1]\lambda)\,!^{\lambda+\mu}=\sum_{s=1}^{\mu(n-1)+1} b_{n,s}m\,!^{s+\mu-1+n\lambda}.$$

Employing the identity

$$\sum_{m=1}^{m} m\,!^{n}=\frac{(m+1)\,!^{n+1}}{n+1}$$

we get the summation

$$(38)\quad \sum_{m=1}^{m} m\,!^{\lambda+\mu}\cdots(m-[n-1]\lambda)\,!^{\lambda+\mu}=\sum_{s=1}^{\mu(n-1)+1} b_{n,s}\frac{(m+1)\,!^{s+\mu+n\lambda}}{s+\mu+n\lambda};$$

the right member may be written as an explicit function of m, n, λ, μ by substituting from (35).

Better summations for the left member of (38) are obtained when either μ is even or n is odd. Thus by means of (11),

$$(39)\quad \sum_{m=1}^{m} m\,!^{\lambda+\mu}\cdots(m-2(n-1)\lambda)\,!^{\lambda+\mu}=\sum_{s=1}^{\mu(n-1)+1} c_{n,s}\frac{(m+s)\,!^{[2n-1]\lambda+2s+\mu-1}}{(2n-1)\lambda+2s+\mu-1};$$

while from (16),

$$(40)\quad \sum_{m=1}^{m} m\,!^{\lambda+2\nu}\cdots(m-[n-1]\lambda)\,!^{\lambda+2\nu}=\sum_{s=1}^{\nu(n-1)+1} d_{n,s}\frac{(m+s)\,!^{n\lambda+2s+2\nu-1}}{n\lambda+2s+2\nu-1};$$

the right members contain approximately but half the number of terms in the right member of (38). Of course (38), (39), (40) can only be spoken of as summations—in a practical sense—when m is large as compared with $n\mu$ (or $n\nu$).

There is a curious connection between (38) and certain Bernoulli polynomials in several variables that I shall consider in detail elsewhere. Limiting ourselves here to the case $\lambda=0$, and modeling after the well known expression of the Bernoulli numbers in terms of the Stirling numbers,* we define ($n_i \geqq 0$)

$$B_n(\xi)=B_{n_1\ldots n_\mu}(\xi_1\cdots\xi_\mu)=\sum_{s=0}^{\infty}\frac{1}{s+1}\sum_{\alpha=0}^{s}(-1)^{\alpha}\binom{s}{\alpha}(\alpha+\xi_1)^{n_1}\cdots(\alpha+\xi_\mu)^{n_\mu}.$$

Note that the outer sum is actually finite and that the inner sum is a generalization of (35) ($\lambda=0$). Then

* J. Worpitzky, "Studien über die Bernoullischen und Eulerschen Zahlen," *Journal für Mathematik*, Vol. 94 (1883), p. 215.

$$\sum_{n} B_n(\xi)\frac{x_1^{n_1}\cdots x_\mu^{n_\mu}}{n_1!\cdots n_\mu!}=\sum_{s=0}^{\infty}\frac{1}{s+1}\sum_{a=0}^{s}(-1)^a\binom{s}{a}e^{(a+\xi_1)x_1+\cdots+(a+\xi_\mu)x_\mu}$$

$$=\sum_{s=0}^{\infty}\frac{1}{s+1}(1-e^{x_1+\cdots+x_\mu})^s\,e^{\xi_1x_1+\cdots+\xi_\mu x_\mu}$$

$$=\frac{(x_1+\cdots+x_\mu)\,e^{\xi_1x_1+\cdots+\xi_\mu x_\mu}}{e^{x_1+\cdots+x_\mu}-1},$$

and this last furnishes a convenient definition for the $B_n(\xi)$. Now on the other hand let

$$S\binom{n}{\xi}=S\begin{pmatrix}n_1\cdots n_\mu\\ \xi_1\cdots\xi_\mu\end{pmatrix}=\sum_{m=0}^{m-1}(m+\xi_1)^{n_1}\cdots(m+\xi_\mu)^{n_\mu}.$$

Evidently then

$$\sum_{n} S\binom{n}{\xi}\frac{x_1^{n_1}\cdots x_\mu^{n_\mu}}{n_1!\cdots n_\mu!}=\sum_{m=0}^{m-1}e^{(m+\xi_1)x_1+\cdots+(m+\xi_\mu)x_\mu}$$

$$=\frac{e^{m(x_1+\cdots+x_\mu)}-1}{x_1+\cdots+x_\mu}\,\frac{(x_1+\cdots+x_\mu)\,e^{\xi_1x_1+\cdots+\xi_\mu x_\mu}}{e^{x_1+\cdots+x_\mu}-1}$$

$$=\sum_{r=1}^{\infty}\frac{(x_1+\cdots+x_\mu)^{r-1}m^r}{r!}\sum_{n}B_n(\xi)\frac{x_1^{n_1}\cdots x_\mu^{n_\mu}}{n_1!\cdots n_\mu!}.$$

Equating coefficients,

$$S\binom{n}{\xi}=\sum_{t_i=0}^{n_i}\binom{n_1}{t_1}\cdots\binom{n_\mu}{t_\mu}B_{n-t}(\xi)\frac{m^{t_1+\cdots+t_\mu+1}}{t_1+\cdots+t_\mu+1}.^*$$

Again if we put

$$T\binom{n}{\xi}=T\begin{pmatrix}n_1\cdots n_\mu\\ \xi_1\cdots\xi_\mu\end{pmatrix}=\sum_{m=0}^{m-1}(-1)^m(m+\xi_1)^{n_1}\cdots(m+\xi_\mu)^{n_\mu},$$

and define the Euler polynomials $E_n(\xi)$ by

$$\frac{2e^{\xi_1x_1+\cdots+\xi_\mu x_\mu}}{e^{x_1+\cdots+x_\mu}+1}=\sum_{n}E_n(\xi)\frac{x_1^{n_1}\cdots x_\mu^{n_\mu}}{n_1!\cdots n_\mu!},$$

it is easily seen that

$$T\binom{n}{\xi}=\tfrac{1}{2}E_n(\xi)-\frac{(-1)^m}{2}E_n(m+\xi).$$

CALIFORNIA INSTITUTE OF TECHNOLOGY,
PASADENA, CALIFORNIA.

* For the function

$$S'\binom{\eta}{\xi}=\sum_{t=0}^{m-1}(d/dt)\,(t+\xi_1)^{n_1}\cdots(t+\xi_\mu)^{n_\mu}$$

there is the much simpler formula

$$S'\binom{\eta}{\xi}=B_n(\xi+m)-B_n(\xi).$$

A CLASS OF DYNAMICAL SYSTEMS ON SURFACES OF REVOLUTION.

By G. Baley Price.

Introduction. A heavy particle on a surface of revolution furnishes the simplest example of a general class of dynamical systems on surfaces of revolution in which the force function depends only on the latitude and not on the longitude. The problem of the heavy particle and other special cases obtained by restricting both the force function and the surface have been treated by Jacobi † and other writers.‡ But even these treatments of special cases of the problem are not in the spirit of modern dynamics and do not use modern methods. The present paper treats the most general case of the problem; the methods used are those of surfaces of section and surface transformations which have been developed by Poincaré and Birkhoff. The problem furnishes a simple example of an integrable dynamical system with two degrees of freedom.

In this exposition only surfaces of genus one are considered, but it is clear that the same methods apply with essentially the same results in the case of surfaces of genus zero. Also, the present problem includes as a special case the determination of the geodesics on the surface.§

1. *The surface and the force function.* In a space with rectangular coördinates (ξ, η, ζ) we shall consider a surface of revolution S which has the ζ-axis as axis of revolution. The surface is generated by rotating about the ζ-axis a simple closed curve whose parametric equations in the (r, ζ) plane of rectangular coördinates are

$$(1) \qquad r = r(x), \qquad \zeta = \zeta(x),$$

where $r(x)$, $\zeta(x)$ are functions satisfying the following hypotheses:

$$(2) \qquad r(x+\omega) = r(x), \qquad \zeta(x+\omega) = \zeta(x),$$

† Jacobi, *Journal für Mathematik*, Vol. 24 (1842), pp. 5-27.

‡ Gustaf Kobb, "Sur le mouvement d'un point matériel sur une surface de révolution," *Acta Mathematica*, Vol. 10 (1887), pp. 89-108; Otto Staude, "Über die Bewegung eines schweren Punktes auf einer Rotationsfläche," *Acta Mathematica*, Vol. 11 (1888), pp. 303-332.

§ B. F. Kimball, "Geodesics on a Toroid," *American Journal of Mathematics*, Vol. 52 (1930), pp. 29-52; G. A. Bliss, "The Geodesic Lines on the Anchor Ring," *Annals of Mathematics*, 2nd ser., Vol. 4 (1902), pp. 1-21.

$$(3) \qquad r(x),\ \zeta(x) \text{ are analytic}, \qquad 0 \leqq x \leqq \omega,$$
$$(4) \qquad r(x) \geqq k > 0, \qquad k \text{ a constant.}$$

We assume also that x is the arc length on the curve (1), i. e.,

$$(5) \qquad r_x^2 + \zeta_x^2 = 1.$$

A subscript x here denotes, as throughout the paper, a derivative with respect to x.

If y is the angle which the plane through the ζ-axis and the point (ξ, η, ζ) forms with the (ξ, ζ) plane, then the parametric equations of S are

$$(6) \qquad \xi = r(x) \cos y, \qquad \eta = r(x) \sin y, \qquad \zeta = \zeta(x).$$

The curves $x =$ constant and $y =$ constant on S will be called *parallels* and *meridians* respectively. It is clear that S is a surface of genus one.

On S we shall consider the motion of a particle of mass m under the action of forces derived from the force function $U = mu(x)$, where $u(x)$ satisfies the following hypotheses:

$$(7) \qquad u(x + \omega) = u(x),$$
$$(8) \qquad u(x) \text{ is analytic}, \qquad 0 \leqq x \leqq \omega.$$

2. *The equations of motion.* Because of (5) we find that the kinetic energy T is given by $T = (m/2)(x'^2 + r^2y'^2)$. Here, as throughout the paper, primes denote derivatives with respect to the time t. The two equations of motion in the Lagrangian form give

$$(9) \qquad r^2y' = c, \qquad \text{[the integral of areas]}$$

where c is a constant of integration, and

$$(10) \qquad x'' = rr_xy'^2 + u_x.$$

The integral of energy is

$$(11) \qquad x'^2 + r^2y'^2 = 2(u + h).$$

We may write (9) and (10) in the form of a first order system of differential equations. In this form all of the usual existence theorems can be applied [R 1, pp. 1-14] (we refer in this fashion to the references at the end).

We may use (9) to eliminate y' from (10) and (11). We obtain

$$(12) \qquad x'' = (c^2r_x + r^3u_x)/r^3,$$
$$(13) \qquad x'^2 = [2r^2(u + h) - c^2]/r^2.$$

For later use we write these equations in the following notation. Define two functions v and w as follows:

(14) $$v = 2r^2(u + h),$$
(15) $$w = -r^3 u_x / r_x, \qquad r_x \neq 0.$$

Then equations (12) and (13) become

(16) $$x'' = r_x(c^2 - w)/r^3, \qquad r_x \neq 0$$
(17) $$x'^2 = (v - c^2)/r^2.$$

There are two further formulas which we shall need later. Let $t(x^0x)$ and $y(x^0x)$ designate the time required for the particle to pass from $x = x^0$ to $x = x$ along a trajectory and the angle y through which it moves respectively. Then from (9) and (17) we obtain the following formulas:

(18)
$$t(x^0x) = \pm \int_{x^0}^{x} \frac{r\,dx}{(v - c^2)^{1/2}};$$
$$y(x^0x) = \pm \int_{x^0}^{x} \frac{c\,dx}{r(v - c^2)^{1/2}},$$

3. *A first surface of section.*

A. The manifold of states of motion and the surface of section. The *manifold of states of motion* [see R 2, Part III for the methods of this section] is the three dimensional manifold in four space obtained by specifying the value of the energy constant h in the integral of energy (11). The integral of areas gives a second integral. Since two integrals exist, the system is said to be *integrable*. The trajectories lie on certain invariant sub-manifolds of (11) which are obtained by specifying the value of c in (9).

According to Birkhoff a *surface of section* is defined as an analytic surface in the manifold of states of motion, "regularly bounded" by a finite number of closed stream lines, cut throughout in the same sense by the stream lines and at least once by every stream line in a fixed interval of time. In this section we propose to obtain a surface of section and show how it can be used to study the trajectories of the system.

In the present section we shall suppose that h has a value such that motion takes place over only a part of S; in section 4 we shall treat the case in which motion takes place over the entire surface. We now make the assumption about the region of motion more precise as follows. Since v is not a constant, we find that it is possible to choose h so that there are regions of motion in which v_x vanishes on only a single parallel. There may be more than one such region for a given value of h, but we consider just one of them. Thus we assume that h is fixed and so chosen that v_x vanishes on only a single parallel, $x = x^*$, in the region of motion under consideration. Then v has a maximum at $x = x^*$, at which we shall assume $v_{xx} < 0$. The motion

on S takes place in the region $v \geqq 0$; this region is bounded by the two parallels on which $v = 0$. Then it is easily found that the parallel $x = x^*$ is a closed periodic trajectory.

We proceed to give a representation of the manifold of states of motion in 3-space. Set

$$\text{(19)} \qquad \tan \psi = ry'/x', \qquad -\pi \leqq \psi \leqq \pi$$

and let ψ vary continuously over the indicated interval as the velocity vector (x', y') turns about (x, y); furthermore, $x' > 0$, $y' = 0$ shall correspond to $\psi = 0$, and the sign of ψ shall be the same as that of y'. Since S is a surface of revolution, ψ is an angle which the trajectory makes at each point with the meridian. To each point (x, y, x', y') in the manifold of states of motion, there corresponds a point (x, y, ψ) in the representation in 3-space; the points (x, y, π) and $(x, y, -\pi)$ are to be considered identical since they correspond to the same point in the original manifold. This representation is valid except at points where $x' = y' = 0$, i. e., except at points on an oval of zero velocity. The representation is certainly valid in the neighborhood of $x = x^*$.

We shall now show that the ring surface

$$x = x^*, \qquad -\pi/2 \leqq \psi \leqq \pi/2, \qquad 0 \leqq y < 2\pi$$

forms a surface of section. In the first place, this surface has two boundaries which correspond to a periodic trajectory traced in the two possible directions. In the second place, we find from the equations of motion that the trajectories oscillate between two parallels on S. Every trajectory therefore cuts through the surface in one and the same sense, and (18) shows that the length of time between any two successive crossings is finite except possibly in the neighborhood of the boundaries of the surface. To determine what happens near the boundaries, we have recourse to the equation of normal displacement for $x = x^*$. If $x = x^* + \epsilon\bar{x}$, $y = y^* + \epsilon\bar{y}$ is a nearby trajectory, a calculation shows that in the limit $\bar{x}$ satisfies the equation

$$\text{(20)} \qquad \frac{d^2\bar{x}}{dt^2} = \left(\frac{v_{xx}}{2r^2}\right)_{x^*} \bar{x}.$$

Hence, since $v_{xx} < 0$ by hypothesis, the limiting length of time between successive crossings as the point of crossing approaches the boundaries of the surface is $2\pi r(2)^{1/2}/(-v_{xx})^{1/2}$. In the third place, no trajectory is tangent to the surface, for a trajectory pierces the surface with the direction components (x', y', ψ'), and x' vanishes only on the boundaries $\psi = \pm\pi/2$. It can be proved easily that the angle at which a trajectory pierces the surface is of the first order in the distance to the boundary. We are thus justified in

calling $x = x^*$, $-\pi/2 \leqq \psi \leqq \pi/2$ a surface of section, because it satisfies the three requirements of the definition.

B. The transformation T. The transformation T is defined as follows: a trajectory which pierces the surface of section at P has its next succeeding intersection at P'. Then $P' = T(P)$.

Now from (19) and (11) we find $ry' = [2(u+h)]^{1/2} \sin\psi$. Then from (9) we have $v^{1/2} \sin\psi = c$. But c is constant along a given trajectory; hence, T has the invariant function $v^{1/2} \sin\psi$. Since x is constant on the surface of section, it follows that *the circles* $\psi =$ *constant are path curves of* T. Furthermore, it is clear from (18) that $y(x^0x)$ is independent of y; hence, T has the form

$$\bar{\psi} = \psi, \qquad \bar{y} = y + \alpha,$$

where α depends on ψ but not on y. Thus T transforms the ring-shaped surface of section into itself by rotating each circle $\psi =$ constant into itself. The amount of rotation α on each circle is called the *rotation number* [R 3, pp. 87-88] of T on that circle. The trajectories oscillate between two parallels on S, and α is the increase in y in one complete oscillation. From (18) we find

$$\alpha = 2 \int_{x_1}^{x_2} \frac{c\,dx}{r(v - c^2)^{1/2}}$$

where x_1 and x_2 are the minimum and maximum values of x respectively on the trajectories for the given value of c. The rotation on the boundaries is found from (20) and (9) to be $2\pi c(2)^{1/2}/r(-v_{xx})^{1/2}$. Thus $\alpha > 0$ when $c > 0$, i. e., when $\psi > 0$, and $\alpha < 0$ when $\psi < 0$. On $\psi = 0$ we have $\alpha = 0$. Also, α is a continuous function of c.

If $\alpha = 2q\pi/p$, where p and q are integers without common factors, on a circle $\psi =$ constant, this circle is rotated into its original position by T^p. Then the circle represents a family of closed, periodic trajectories which we designate as of type (p, q). Here $p > 0$ represents the number of complete oscillations between two parallels which the trajectory makes before closing, and q, positive or negative, is the number of multiples of 2π by which y increases.

Let α_M be the maximum rotation number on the entire surface of section and α_B the rotation number on $\psi = \pi/2$. Then since α varies continuously on the circles $\psi =$ constant, we see at once that the following theorem is true.

THEOREM. *If* $\alpha_B < \alpha_M$, *there are at least two families of closed, periodic trajectories of each type* (p, q), *where* $\alpha_B < |2q\pi/p| < \alpha_M$; *in any case,*

there is at least one family of closed, periodic trajectories of each type (p, q) *where* $0 \leqq |2q\pi/p| < \alpha_B$.

We state without going into details at this point that the circles on which α is an irrational multiple of 2π represent families of recurrent trajectories.

It is possible to show that there is an invariant volume integral in the manifold of states of motion, and that T has an invariant area integral on the surface of section. The surface transformation here considered is then a degenerate case of Poincaré's Last Geometric Theorem [R 4; R 2, p. 294]. It illustrates also one of the three general types of fixed points in surface transformations [R 3, p. 4, type III]. The surface of section and transformation could be extended to some extent to other cases of the problem, but we prefer to consider now a second type of surface of section, which is more interesting in the general case.

4. *A second surface of section.*

A. The manifold of states of motion. On the surface S consider a general dynamical system in which the force function is $U = mu(x, y)$, where $u(x, y)$ is an arbitrary analytic function which is periodic with periods ω and 2π in x and y respectively. Then the equations of motion are

$$(21) \qquad x'' = u_x + rr_x y'^2, \qquad y'' = (u_y - 2rr_x x' y')/r^2,$$

and the integral of energy is

$$(22) \qquad x'^2 + r^2 y'^2 = 2(u + h).$$

According to the definition given in section 3, the equation of the manifold of states of motion is (22). We can give a representation of this manifold in 3-space as follows. Set

$$(23) \qquad \tan\phi = x'/ry', \qquad -\pi \leqq \phi \leqq \pi$$

and let ϕ vary continuously over the interval indicated as the vector (x', y') turns about (x, y). Also, let $\phi = 0$ when $x' = 0$, $y' > 0$, and let ϕ have the same sign as x'. Since S is a surface of revolution, ϕ is an angle which the trajectory makes with the parallel at each point. Corresponding to each state of motion (x, y, x', y') there is a point (x, y, ϕ) in the representation; the points (x, y, π) and $(x, y, -\pi)$ are to be considered identical. The totality of points (x, y, ϕ) form a manifold M in 3-space. This representation is valid except at points where $x' = y' = 0$.

Now the totality of states of motion along a trajectory corresponds to a curve, a *stream line*, in M. The differential equations of these stream lines can be found as follows. From (23), $x' = \rho \sin\phi$, $ry' = \rho\cos\phi$. Substitute

in (22) to determine ρ. A straightforward calculation gives ϕ'. The results are

$$(24)\qquad \begin{cases} x' = X(x, y, \phi) \equiv [2(u+h)]^{1/2} \sin\phi \\ y' = Y(x, y, \phi) \equiv (1/r)[2(u+h)]^{1/2} \cos\phi \\ \phi' = \Phi(x, y, \phi) \equiv \dfrac{ru_x \cos\phi - u_y \sin\phi + 2r_x(u+h)\cos\phi}{[2r^2(u+h)]^{1/2}} \end{cases}$$

Now the time t can be eliminated completely from (24); we obtain

$$(25)\qquad dx/X = dy/Y = d\phi/\Phi.$$

The trajectories thus appear in M as the stream lines of the steady fluid motion defined by (25).

A fundamental property of the steady flow in M is stated in the following lemma.

LEMMA 1. *The volume integral*

$$(26)\qquad \iiint r(x)\, dx\, dy\, d\phi$$

is invariant in the steady flow in M defined by (25).

A necessary and sufficient condition that (26) be invariant is that [R 5, pp. 285-286; R 6, vol. III, pp. 1-6]

$$\frac{\partial(rX)}{\partial x} + \frac{\partial(rY)}{\partial y} + \frac{\partial(r\Phi)}{\partial\phi} \equiv 0$$

and we find by substituting from (24) that this condition is satisfied.

We may think of $r(x)$ as the density function in M. Then (26) represents the mass.

B. The surface of section. We return now to a consideration of the special systems treated in this paper, i. e., systems in which $u = u(x)$. We shall assume that h is so chosen that $v > 0$ for all values of x. Then the meridian $y = 0$ is a closed periodic trajectory, and we shall show how a surface of section of a certain type can be constructed from it. In the manifold M, a ring-shaped surface R is defined by $y = 0$, $y' \geqq 0$.

Now in the first place, the boundaries of R are $y = 0$, $y' = 0$ with either $x' > 0$ or $x' < 0$. They are therefore two closed stream lines which correspond to a closed trajectory traced in the two possible directions. In the second place, no stream line is tangent to R, and the angle at which a stream line cuts R is of the first order in the distance to the boundary, for if θ is the angle,

$$\sin\theta = y'/(x'^2 + y'^2 + \phi'^2)^{1/2}.$$

Now the distance to the boundaries is $|\pm\pi/2-\phi|$. Then

$$\lim_{\phi\to\pm\pi/2}\frac{\theta}{|\pm\pi/2-\phi|}=\lim_{\phi\to\pm\pi/2}\frac{\sin\theta}{|\pm\pi/2-\phi|}.$$

Substitute now for $\sin\theta$, and then substitute from the equations corresponding to (24). We find that the limit is $1/r\neq 0$, from which the stated results follow.

In the third place, we must consider the intersections of the trajectories with R. Now we may consider $y'\geqq 0$ on every trajectory, because negative values of y' merely give the same trajectories traced in the opposite direction. The trajectory $y=0$, $y'=0$ forms the boundaries of R, but no other trajectory with $y'=0$ has any point in common with R. Therefore R does not satisfy the requirement for a surface of section that *all* trajectories intersect it. Now (9) shows that all trajectories which are not meridians have $y'>0$ always; hence, sooner or later they cross R. However, as c approaches zero, the maximum value of y' approaches zero; hence, the interval of time between successive crossings of R becomes infinite. The intersections with R approach the boundaries as c approaches zero. In any closed region $r^2y'\geqq c^0>0$ inside the ring R the interval of time between successive crossings is uniformly bounded. We see therefore that R fails to satisfy a second requirement for a surface of section.

LEMMA 2. *In the sense explained, the surface* $R: y=0,\ y'\geqq 0$ *forms a surface of section.*

C. The transformation T. The transformation T on R is defined in the usual way: a stream line which crosses R at P has its first succeeding crossing at P'. Then $P'=T(P)$. This transformation carries R into itself in a one-to-one and continuous manner and is analytic in the interior of the ring. Results to be established presently make it clear that T cannot be analytic along the boundaries of R.

A fundamental property of T is established in the following lemma.

LEMMA 3. *The transformation T on R has the positive invariant area integral*

$$\iint [2(u+h)]^{1/2}\cos\phi\, dx\, d\phi, \tag{27}$$

and its value over the entire ring is finite.

Consider a region σ on R and the region σ' into which it is transformed by T. By lemma 1, the mass of the tube of stream lines bounded by σ and σ' is invariant as it moves along. The rate of decrease of mass at the base σ is [R 5, pp. 286-287]

$$\int\int_{\sigma} ry'\,dx\,d\phi,$$

and the rate of increase at the base σ' is given by the same integral extended over the region σ'. It follows that these two integrals are equal; hence, substituting for y' from (24) we find that (27) is an invariant integral. The value of the integral extended over the entire ring is $2\int_0^{\omega}[2(u+h)]^{1/2}\,dx$, and this is positive and finite.

If we consider $[2(u+h)]^{1/2}\cos\phi$ the density of R, then (27) is the mass.

The following lemma gives a second important property of T.

LEMMA 4. *The function $F \equiv v\cos^2\phi$ is invariant under T.*

From (9) and the equations corresponding to (24) we find that $F = c^2$. But since c is constant along a given trajectory, the lemma follows.

The existence of F is a direct consequence of the fact that the system is integrable, i. e., that two integrals (9) and (11) are known. Now F is positive over the interior of R and vanishes on the boundaries. Furthermore, it is symmetric in $\phi = 0$, and for a fixed value of x decreases monotonically as $|\phi|$ approaches $\pi/2$. It follows that the critical points of F, i. e., points (x, ϕ) at which the two first partial derivatives of F vanish [Morse's definition], which are interior points of R lie on the line $\phi = 0$. The level curves of F are path curves of T.

D. The path curves of T. As we have just seen, the value of c determines the nature of the path curve $F = c^2$ which corresponds to the trajectories for the given value of c. First we shall determine the types of path curves, and then study the nature of the corresponding trajectories.

We shall suppose that R is taken as a ring bounded by two concentric circles. The coördinates on R are (x, ϕ), where x varies from 0 to ω around the ring in the counter clockwise direction, and ϕ varies from $-\pi/2$ on the inner boundary to $\pi/2$ on the outer boundary. The path curves $F = c^2$ are symmetric in the circle $\phi = 0$.

Now plot $z = F$ over the ring R, and plot also $z = c^2$, a plane parallel to the ring. The projections on R of the curves of intersection of $z = F$ and the plane are the path curves. By letting c^2 vary from 0 to the maximum value of v, we obtain the totality of path curves.

In the first place, $z = F$ has a certain number of critical points, which are of special importance. For convenience we agree that *critical points of F* shall include only points on $\phi = 0$. Then the critical points of F are those

points and only those points on $\phi = 0$ for which $v_x = 0$. Now a calculation shows that v_x is given by

$$(28) \qquad v_x = 2[2rr_x(u+h) + r^2 u_x],$$

$$(29) \qquad v_x = 2r_x(v - w)/r, \qquad r_x \neq 0.$$

We now find from the equations of motion (12), (13) and (16), (17) that a necessary and sufficient condition that the parallel $x = x^*$ be a trajectory is that v_x vanish for $x = x^*$. Furthermore, there are two types of vanishing of v_x. From (28) the first occurs when $r_x = u_x = 0$, and from (29) the second when $v - w = 0$, $r_x \neq 0$. We may thus distinguish two types of critical points of $z = F$. The first of these two does not vary with h, but since w is independent of h and v depends on h, the second type of critical point does vary with h. Finally, since critical points of $z = F$ correspond to trajectories which are parallels, we see that *critical points of F are invariant points of T.*

Now for certain values of c^2, the plane $z = c^2$ passes through critical points of $z = F$ which are not maxima (critical points which correspond to minima or points of inflection of the function $z = v$). Let these values of c^2, in the order of increasing magnitude, be designated by $c_0{}^2, c_1{}^2, \cdots, c_k{}^2$. Then each of the path curves $F = c_i{}^2$, $(i = 0, 1, \cdots, k)$, passes through an invariant point of T. If $z = v$ has a minimum at the critical point, the path curve has two branches which pass through the invariant point. If $z = v$ has a point of inflection, the path curve has a cusp at the critical point. We designate by P a critical point of $z = F$, and by C_1 a path curve, exclusive of the critical point or points, of the set $F = c_i{}^2$, $(i = 0, 1, \cdots, k)$.

For $c^2 = 0$, the trajectories are meridians. The corresponding curves $F = 0$ on R are the two boundaries. Now $F = c^2$ gives two path curves, designated by C_2, when $0 < c^2 < c_0{}^2$, one of which lies in the region between $F = c_0{}^2$ and the outer boundary of R, and the other of which lies between the same curve and the inner boundary. They are simple closed curves which can be deformed through points of R into the two boundaries of R.

Finally, $F = c^2$ gives one or more simple closed curves when $c^2 > c_0{}^2$, $c^2 \neq c_i{}^2$, $(i = 0, 1, \cdots, k)$. We designate these curves by C_3; they fill out the regions between the curves C_1.

We have thus found that the system has five types of trajectories: parallel trajectories corresponding to critical points P; meridian trajectories corresponding to the boundaries of R; and three other types corresponding to the path curves C_1, C_2, C_3.

E. Further properties of T. In the regions of R occupied by the path curves C_2 and C_3, it is possible to use x and c as coördinates instead of x

and ϕ. The transformation to the new coördinates is accomplished by means of the relation $c = v^{1/2} \cos \phi$. We thus find that the invariant area integral (27) expressed in terms of the new coördinates (x, c) is

$$\int\int \frac{c\, dx\, dc}{r(v - c^2)^{1/2}}. \tag{30}$$

Since (30) is invariant, it follows that the integral

$$\int \frac{c\, dx}{r(v - c^2)^{1/2}} \tag{31}$$

is invariant under T on the path curves $F = c^2$ [R 3, pp. 93-94]. This integral diverges of course if integrated along a path curve C_1 up to a critical point of F.

Consider the path curves C_2. First we replace x by a new coördinate τ as follows. Set

$$\frac{\tau}{2\pi} \int_0^\omega \frac{c\, dx}{r(v - c^2)^{1/2}} = \int_0^x \frac{c\, dx}{r(v - c^2)^{1/2}}. \tag{32}$$

Then as x varies from 0 to ω, τ varies from 0 to 2π. We may obviously think of (c, τ) as polar coördinates.

LEMMA 5. *In the coördinates (c, τ), T is the rotation $c' = c$, $\tau' = \tau + \alpha$ on the path curves C_2.*

Let $x_1 : \tau_1$ and $x_2 : \tau_2$ be two points on a curve C_2 which are transformed by T into $x'_1 : \tau'_1$ and $x'_2 : \tau'_2$. Then we find

$$\frac{\tau_2 - \tau_1}{2\pi} \int_0^\omega \frac{c\, dx}{r(v - c^2)^{1/2}} = \int_{x_1}^{x_2} \frac{c\, dx}{r(v - c^2)^{1/2}},$$

and a similar equation in which x_1, x_2, τ_1, τ_2 are primed. But since (31) is invariant under T, we have $\tau'_2 - \tau'_1 = \tau_2 - \tau_1$. Set $\tau'_1 - \tau_1 = \alpha$ and drop the subscript 2. Then $\tau' = \tau + \alpha$ and the lemma is proved.

The number α is the rotation number of T on the given path curve. We compute α as follows.

$$\frac{\alpha}{2\pi} \int_0^\omega \frac{c\, dx}{r(v - c^2)^{1/2}} = \int_{x_1}^{x'_1} \frac{c\, dx}{r(v - c^2)^{1/2}}.$$

Now from (18)

$$\pm \int_{x_1}^{x'_1} \frac{c\, dx}{r(v - c^2)^{1/2}} = 2\pi$$

the plus or minus sign being used according as the path curve lies in the region $\phi > 0$ or $\phi < 0$. Substituting from the second equation in the first, we find

$$\alpha = \pm 4\pi^2 \Big/ \int_0^\omega \frac{c\, dx}{r(v - c^2)^{1/2}}. \tag{33}$$

From this equation we see that on the path curves C_2 in the region $\phi > 0[\phi < 0] 0 < \alpha < +\infty[-\infty < \alpha < 0]$; as c decreases from c_0 to 0, α increases [decreases] monotonically and continuously from 0 to $+\infty$ [0 to $-\infty$]. This statement proves the following lemma.

LEMMA 6. *On the path curves C_2, α takes on every value except $\alpha = 0$ once and only once.*

Now consider the path curves C_3. We introduce a parameter τ on these curves as follows. Let x_1 and x_2 be the minimum and maximum values respectively of x on a given curve C_3. Set

$$\frac{\tau}{\pi}\int_{x_1}^{x_2}\frac{c\,dx}{r(v-c^2)^{1/2}} = \pm\int_{x_1}^{x}\frac{c\,dx}{r(v-c^2)^{1/2}}.$$

The integral on the right is to be integrated around the path curve in the counter clockwise direction, the plus sign being taken when $\phi > 0$ and the minus sign when $\phi < 0$. Then as the point with abscissa x traces the curve, τ increases from 0 to 2π. We may think of (c, τ) as polar coördinates on the curves C_3.

The proof of the following lemma is similar to that of lemma 5.

LEMMA 7. *In the coördinates (c, τ), T is the rotation $c' = c$, $\tau' = \tau + \alpha$ on the curves C_3.*

From (18) we find that on the curves C_3, α is given by

$$\alpha = 2\pi^2 / \int_{x_1}^{x_2}\frac{c\,dx}{r(v-c^2)^{1/2}}. \tag{34}$$

The rotation number on every path curve C_3 is therefore positive; α is a continuous function of c, and α approaches zero as c approaches c_i, $(i = 0, 1, \cdots, k)$.

The path curves C_3 fill out regions of two kinds: (a) ring shaped regions bounded on the inside and outside by a path curve C_1; (b) circular regions about an invariant point P, bounded on the outside by a curve C_1. Let α_M denote the maximum rotation number in these regions, and α_P the rotation number at a point P.

LEMMA 8. *On the curves C_3 in a ring region, α takes on each value on the interval $0 < \alpha < \alpha_M$ at least twice; on the curves C_3 in a circular region, α takes on each value $0 < \alpha < \alpha_P$ at least once, and if $\alpha_M > \alpha_P$, it takes on each value on the interval $\alpha_P < \alpha < \alpha_M$ at least twice.*

Now the curves C_1 are composed of arcs terminated by critical points of F. On each of these arcs introduce a new parameter τ as follows. Choose an interior point x^0 of the arc and set

$$\tau = \pm \int_{x^0}^{x} \frac{c\,dx}{r(v-c^2)^{1/2}},$$

where the plus (minus) sign is used when $\phi > 0$ ($\phi < 0$). Then τ increases from $-\infty$ to $+\infty$ as x increases (decreases) from one end of the arc to the other in $\phi > 0$ ($\phi < 0$). Let the points $x : \tau$ and $x^0 : 0$ be carried into $x' : \tau'$ and $x^* : h$ by T. Then since (31) is invariant under T, we find $\tau = \tau' - h$ or $\tau' = \tau + h$. From (18) we find

$$\int_{x^0}^{x^*} \frac{c\,dx}{r(v-c^2)^{1/2}} = 2\pi;$$

and hence T is the translation $\tau' = \tau + 2\pi$.

LEMMA 9. *In terms of the parameter τ, T on each arc of a curve C_1 is the translation $\tau' = \tau + 2\pi$; with reference to R, the motion is counter clockwise or clockwise according as $\phi > 0$ or $\phi < 0$.*

We have thus determined completely the nature of the transformation on R. We see that *the transformation T on the ring R is such that points are advanced in the region $\phi > 0$ and regressed in the region $\phi < 0$.* In paragraph D it was shown by using the equations of motion directly that the critical points of F are invariant points of T; this fact follows also from the properties of T which have been established in the present paragraph. We now see that T satisfies all the hypotheses of Poincaré's Last Geometric Theorem [R 4], for it is one-to-one and continuous in R; it has an invariant area integral; and it advances points in the neighborhood of the boundary $\phi = \pi/2$ and regresses points in the neighborhood of the boundary $\phi = -\pi/2$. It follows that T has at least two invariant points.

In the present problem it has been shown already that the invariant points of T are the critical points of F. At this point it is impossible to avoid the notion of multiple invariant points [R 2, p. 286]. It has been seen that each zero of v_x gives an invariant point; also there is one type of zero of v_x which varies with the energy constant h. Then when two zeros of v_x combine to form a multiple zero, two invariant points unite to form what it is natural to call a multiple invariant point. In general we may agree that an invariant point is multiple in the same sense in which the zeros of v_x are multiple. It can be shown that an invariant point which is multiple in the present sense is also multiple according to the definition of Birkhoff.

Let $x = x^*$ be a parallel trajectory and c^* the corresponding value of c. We can expand $F - c^{*2}$ in a Taylor's series about $(x^*, 0)$ and obtain

$$(35) \qquad v\cos^2\phi - c^{*2} = \tfrac{1}{2}\left[v_{xx}(x^*)(x-x^*)^2 - 2v(x^*)\phi^2\right] + \cdots.$$

Now suppose that v has a minimum at $x = x^*$ with $v_{xx}(x^*) > 0$. Then

(35) shows that the path curve $F - c^{*2} = 0$ has two branches through $(x^*, 0)$ with distinct tangents there. From lemma 9 we see that T moves points toward the invariant point on one of the branches, and away from the invariant point on the other. Thus *an invariant point with* $v_{xx}(x^*) > 0$ *is unstable and of hyperbolic type*, and from lemma 9 we see at once that it is *directly unstable* [R 2, pp. 286-287; also R 3].

Now suppose that v has a maximum at $x = x^*$ with $v_{xx}(x^*) < 0$. Then (35) shows that the corresponding invariant point is surrounded by path curves of type C_3. Then *the invariant point is stable and of elliptic type.*

Finally, *the invariant points at which* $v_{xx} = 0$ *are multiple.* If v has a minimum at $x = x^*$, there are two branches of a path curve C_1 through the invariant point $(x^*, 0)$, but the two branches now have a common tangent. The invariant point is unstable. If v has a maximum at $x = x^*$, the invariant point is surrounded by path curves C_3 and is stable. If v has a point of inflection with a horizontal tangent, the path curve through the corresponding invariant point has a cusp there; the invariant point is unstable.

These results on stability may be compared with those yielded by equation (20).

F. Asymptotic, periodic, and recurrent trajectories. We shall now make use of the properties of T to determine the nature of the trajectories of the system. In the first place, the curves C_1 are made up of arcs terminated by invariant points. On each arc T is a translation toward one of the invariant points in one sense of the time and toward the other (which may be identical with the first) in the opposite sense. The invariant points represent parallel trajectories. The other points of the arc represent trajectories which are asymptotic to these parallels in the two senses of the time.

From lemma 5 we see that the path curves C_2 correspond to trajectories on which neither x' nor y' ever changes sign, i. e., the trajectories wind constantly about S in one direction. On the other hand, the path curves C_3 correspond to trajectories which oscillate between two parallels on S as we see from lemma 7. The two parallels correspond to the maximum and minimum values of x on the path curve. In this connection it is to be remembered that each path curve is merely a section of a cylindrical manifold in the manifold of states of motion on which the stream lines lie.

We shall now consider the periodic trajectories other than the parallel trajectories. On the path curves C_2 and C_3 the transformation T is a rotation $\tau' = \tau + \alpha$. Now if α has the form $2p\pi/q$, where p, q are integers without common factors, then T^q rotates the curve through p complete revolutions,

and every point is transformed into itself. The points on such curves then represent closed, periodic trajectories. These trajectories will be said to be of type $k: (p, q)$ where $k = 2$ or 3 is the subscript of the corresponding path curve. The numbers have the following geometric significance. The number $q > 0$ is the number of multiples of 2π by which y increases along the trajectory before it closes. For a trajectory of type $2: (p, q)$, $p > 0$ $(p < 0)$ is the number of multiples of ω by which x increases (decreases) along the trajectory; for a trajectory of type $3: (p, q)$, $p > 0$ is the number of complete oscillations between two parallels. Furthermore, it is a natural extension to consider the meridians as trajectories of type $2: (p, 0)$ for p any positive or negative integer not zero.

THEOREM 1. *Through each point of S there passes one and only one closed periodic trajectory of type $2: (p, q)$ where $p \neq 0$, $q \geqq 0$.*

If the theorem is true for the points of S on $y = 0$, it is true for all the points of S. The trajectories of type $2: (p, q)$ through points on $y = 0$ are represented on R by the points of the path curves C_2. Now by lemma 6 there is one and only one curve C_2 on which α takes on any given value, not zero; the existence of one and only one trajectory of type $2: (p, q)$ with $q > 0$ is thus established. But the meridian $y = 0$ corresponds to the trajectories with $q = 0$. The theorem is proved. In this connection compare Birkhoff's existence theorem for closed trajectories of minimum type [R 2, p. 219].

The trajectories which correspond to the points on a path curve C_1, C_2, or C_3 will be called a *family of trajectories*. From lemma 8 we obtain at once the following theorem.

THEOREM 2. *Represented among the path curves C_3 which form a ring region there are at least two families of closed periodic trajectories of each type $3: (p, q)$ where $0 < 2p\pi/q < \alpha_M$. Represented among the path curves C_3 which form a circular region there is at least one family of closed periodic trajectories of each type $3: (p, q)$ where $0 < 2p\pi/q < \alpha_P$, and if $\alpha_M > \alpha_P$, at least two of each type $3: (p, q)$ where $\alpha_P < 2p\pi/q < \alpha_M$.*

There remain to be considered only the trajectories represented by path curves C_2 and C_3 on which α is an irrational multiple of 2π. We shall now show that these trajectories are of the type known as *recurrent*. Instead of going into the details of defining recurrent trajectories, we state the following theorem which gives their characteristic property [R 7, p. 312]: *The necessary and sufficient condition that a stable trajectory be recurrent is that given any $\epsilon > 0$ it is possible to find a t^* such that in any time interval of length t^* the trajectory comes within a distance ϵ of every point of the entire*

trajectory. In this theorem, as well as in the following one, the trajectory is considered as a stream line in (x, y, ϕ) space. Every trajectory of our system is stable in the sense in which the term is used in the theorem.

THEOREM 3. *The path curves C_2 and C_3 on which α is an irrational multiple of 2π represent recurrent trajectories.*

Consider two trajectories with their initial states of motion represented by two points τ_1 and τ_2 on a curve C_2 or C_3, on C^*_2 say. Now since the trajectories vary analytically with the initial conditions, it is possible to find an η such that when $|\tau_2 - \tau_1| < \eta$, the distance between the points of the trajectories for corresponding values of t is less than ϵ, at least until the next succeeding intersection of the trajectories with the surface of section. Let τ_0 be any point of C^*_2 and $\tau_1, \tau_2, \cdots$ its transforms under $T, T^2, \cdots$. Because of the nature of α, it is possible to find an N such that $|\tau - \tau_j| < \eta$, where τ is any point of C^*_2 and τ_j is some point of $\tau_0, \tau_1, \cdots, \tau_N$. Now it is possible to find a t^* such that any trajectory represented by a point of C^*_2 crosses the surface of section N times in any time interval of length t^*. It follows that any trajectory represented by a point of C^*_2 comes within a distance ϵ of every point of all trajectories represented by points on C^*_2 in any interval of time of length t^*. Hence, the trajectory is recurrent, and the proof is complete.

As remarked by Birkhoff, the recurrent motions of an integrable dynamical system are of the type known as *continuous*. Morse has treated *discontinuous* recurrent trajectories [R 8 and 9].

HARVARD UNIVERSITY.

REFERENCES.

1. Birkhoff, "Dynamical Systems," *American Mathematical Society Colloquium Publications* No. 9.

2. Birkhoff, "Dynamical Systems with Two Degrees of Freedom," *Transactions of the American Mathematical Society*, Vol. 18 (1917), pp. 199-300.

3. Birkhoff, "Surface Transformations and Their Dynamical Applications," *Acta Mathematica*, Vol. 43 (1922), pp. 1-119.

4. Birkhoff, "Proof of Poincaré's Geometric Theorem," *Transactions of the American Mathematical Society*, Vol. 14 (1913), pp. 14-22.

5. Appell, *Traité de Mécanique Rationnelle*, Vol. 3 (1921).

6. Poincaré, *Méthodes Nouvelles de la Mécanique Céleste*, Vols. 1, 3.

7. Birkhoff, "Quelques Théorèmes sur le Mouvement des Systèmes Dynamiques," *Bulletin de la Société Mathématique de France*, Vol. 40 (1912), pp. 305-323.

8. Morse, "A One-to-One Representation of Geodesics on a Surface of Negative Curvature," *American Journal of Mathematics*, Vol. 43 (1921), pp. 33-51.

9. Morse, "Recurrent Geodesics on a Surface of Negative Curvature," *Transactions of the American Mathematical Society*, Vol. 22 (1921), pp. 84-100.

A BOUNDARY VALUE PROBLEM ASSOCIATED WITH THE CALCULUS OF VARIATIONS.†

By William T. Reid.

1. Introduction. Let η denote a set of variables $(\eta_1, \eta_2, \cdots, \eta_n)$ each of which is a function of the real variable x and denote the end values of these functions at x_1 and x_2 by $\eta(x_1)$ and $\eta(x_2)$. Let $\omega(x, \eta, \eta')$ be for each x on the interval x_1x_2 a homogeneous quadratic form in the variables η_i, η'_i, and denote by $H[\eta(x_1), \eta(x_2)]$ and $G[\eta(x_1), \eta(x_2)]$ two homogeneous quadratic forms in the variables $\eta_i(x_1)$, $\eta_i(x_2)$. Now consider the problem of minimizing the expression

$$I[\eta] = 2H[\eta(x_1), \eta(x_2)] + \int_{x_1}^{x_2} 2\omega(x, \eta, \eta')\, dx \tag{1.1}$$

in the class of arcs

$$\eta_i = \eta_i(x) \qquad x_1 \leqq x \leqq x_2 \qquad (i = 1, 2, \cdots, n) \tag{1.2}$$

which satisfy a set of ordinary linear differential equations of the first order

$$\Phi_\alpha(x, \eta, \eta') = 0 \qquad (\alpha = 1, 2, \cdots, m < n), \tag{1.3}$$

which satisfy end-conditions

$$\Psi_\gamma[\eta(x_1), \eta(x_2)] = 0 \qquad (\gamma = 1, 2, \cdots, p \leqq 2n), \tag{1.4}$$

where the functions Ψ_γ are homogeneous of the first degree in their arguments, and which give a fixed value to the expression

$$2G[\eta(x_1), \eta(x_2)] + \int_{x_1}^{x_2} \eta_i(x) K_{ij}(x) \eta_j(x)\, dx \tag{1.5}$$

This problem of the calculus of variations may be reduced to a problem of Mayer of the type considered by Bliss.‡ The boundary value problem consisting of the Euler-Lagrange equations and the transversality conditions for this problem of the calculus of variations will be treated in the present paper.

In § 2 will be stated the hypotheses upon which the analysis is based, and in §§ 3 and 4 some properties of the boundary value problem are dis-

† The present paper is the revised form of a paper written while the author was a National Research Fellow, and which was presented to the American Mathematical Society, September 9, 1931.

‡ G. A. Bliss, *Transactions of the American Mathematical Society*, Vol. 19 (1918), pp. 305-314.

cussed. In § 5 there are defined the successive classes S_i $(i = 1, -1, 2, -2, \cdots)$ of arcs $\eta(x)$ in which we shall consider the problem of minimizing $I[\eta]$. It is proven that if the class S_i is not empty then the greatest lower bound of $I[\eta]$ in this class is the absolute value of a characteristic number for the boundary value problem. In § 6 there are stated sufficient conditions for the boundary value problem to have infinitely many characteristic numbers. Finally, § 7 is devoted to some expansion theorems in terms of the characteristic solutions of our boundary value problem for the special case when the coefficients of $G[\eta(x_1), \eta(x_2)]$ are zero.

The problem defined above includes as special cases many boundary value problems which have previously been considered. If $\| K_{ij}(x) \|$ is the unit matrix and the coefficients of $H[\eta(x_1), \eta(x_2)]$ and $\Phi_a(x, \eta, \eta')$ are zero, while the end-conditions (1.4) reduce to

$$\eta_i(x_1) = 0 = \eta_i(x_2) \qquad (i = 1, 2, \cdots, n),$$

the boundary value problem is of the type considered by Hickson.† The proof of the existence of successive characteristic numbers for our boundary value problem in § 5 is by the direct methods of the calculus of variations and closely parallels the method used by Mason ‡ in treating second order differential equations. The method, therefore, is quite different from that used by Hickson in treating the special problem stated above. It should be mentioned, however, that if the differential system satisfies certain normality conditions which are more stringent than those imposed here, then the proof of the existence of successive characteristic numbers may be made in a manner entirely analogous to that used by Hickson.

For the general problem of Bolza with variable end-points the second variation is expressible in the form (1.1), where the variables η are the variations of a one-parameter family of admissible arcs, together with the variations of the end-points of such a family.§ Morse ¶ has obtained sufficient conditions for the problem of Bolza when the non-tangency condition is satisfied, in which case the variations of the end-points of the family may be expressed in terms of the variations of the functions of the family. He has defined for this problem of Bolza the accessory boundary value problem

† A. O. Hickson, *Transactions of the American Mathematical Society*, Vol. 31 (1929), pp. 563-579.

‡ M. Mason, *The New Haven Colloquium of the American Mathematical Society*, pp. 173-222; in particular, pp. 207-222.

§ T. F. Cope, *Dissertation*, University of Chicago, 1927.

¶ M. Morse, *American Journal of Mathematics*, Vol. 53 (1931), pp. 517-546. This paper will be referred to as M.

as the above defined problem with $\| K_{ij} \|$ the identity matrix and the coefficients of $G[\eta(x_1), \eta(x_2)]$ all zero. The condition that the smallest characteristic number of this boundary value problem be positive plays in the sufficiency theorem for the problem of Bolza considered by Morse the same rôle that the requirement that the Jacobi condition be satisfied in the strong sense plays in sufficiency theorems for the simpler problems of the calculus of variations.† Morse has shown that there exist infinitely many characteristic numbers for this accessory boundary value problem. The proof of this existence theorem is quite different from the proof of the existence theorem for the general boundary value problem here discussed, and depends upon the results obtained by Morse on the calculus of variations in the large.‡

2. Notation and preliminary remarks. In the later discussion the following subscripts will have the ranges indicated: $i, j, k = 1, 2, \cdots, n$; $\alpha = 1, 2, \cdots, m$; $\pi, \kappa, \tau = 1, 2, \cdots, 2n$; $\gamma = 1, 2, \cdots, p$; $\theta = 1, 2, \cdots, 2n - p$. The repetition of a subscript in a single term of an expression will denote summation with respect to that subscript over the range on which it is defined. Partial derivatives of ω and Φ_α with respect to the variables η_i and η'_i will be denoted by writing the variable as a subscript; correspondingly, the derivatives of the functions Ψ_γ, H and G with respect to $\eta_i(x_1)$ and $\eta_i(x_2)$ will be denoted by the subscripts η_{i1} and η_{i2} respectively.

An arc $\eta \equiv [\eta_i(x)]$ will be called *differentially admissible* if the functions $\eta_i(x)$ are of class D' § on x_1x_2 and satisfy (1. 3) on this interval. An arc which satisfies the end-conditions (1. 4) will be said to be *terminally admissible*. Finally, an arc which is both differentially and terminally admissible will be called *admissible*.¶

The analysis is based upon the following hypotheses:

(H1) All the coefficients in (1. 1), (1. 3), (1. 4) and (1. 5) are real, and the functions $\omega_{\eta'_i\eta'_j}$, $\omega_{\eta'_i\eta_j}$, $\Phi_{\alpha\eta'_i}$ are of class C', while the functions

† Recently Bliss has given sufficient conditions for the problem of Bolza for the more general case in which the non-tangency condition is not assumed to be satisfied. See G. A. Bliss, *Annals of Mathematics*, Vol. 33 (1932), pp. 261-274. Instead of stating the analogue of the Jacobi condition in terms of the smallest characteristic number of a boundary value problem involving differential equations and end-conditions, Bliss has replaced this boundary value problem by a corresponding algebraic problem.

‡ M. Morse, *Transactions of the American Mathematical Society*, Vol. 31 (1929), pp. 379-404.

§ We shall use the terminology for classes of functions introduced by Bolza. See O. Bolza, *Lectures on the Calculus of Variations*, 1904, p. 7.

¶ This terminology is due to Morse. See M, p. 518.

$\omega_{\eta_i\eta_j}$, $\Phi_{a\eta_j}$ and K_{ij} are continuous on x_1x_2. The matrices $\| \omega_{\eta'_i\eta'_j} \|$, $\| \omega_{\eta_i\eta_j} \|$ and $\| K_{ij} \|$ are symmetric and $\| \Phi_{a\eta'_j} \|$ is of rank m at each point on x_1x_2, while the constant matrix $\| \Psi_{\gamma\eta_{i1}}; \Psi_{\gamma\eta_{i2}} \|$ has rank p.

(H2) The symmetric matrix

$$\left\| \begin{matrix} \omega_{\eta'_i\eta'_j} & \Phi_{a\eta'_i} \\ \Phi_{a\eta'_j} & 0 \end{matrix} \right\| \tag{2.1}$$

is non-singular on x_1x_2.

If η is a normal minimizing arc for the problem of the calculus of variations defined in § 1, it follows that there exists a constant λ and functions $\mu_a(x)$ such that if we define

$$\Omega(x, \eta, \eta', \mu) = \omega(x, \eta, \eta') + \mu_a(x)\Phi_a(x, \eta, \eta') \tag{2.2}$$

and

$$J_i(\eta, \mu) = d\Omega_{\eta'_i}/dx - \Omega_{\eta_i}, \tag{2.3}$$

then on every sub-arc between corners of η the differential equations

$$J_i(\eta, \mu) + \lambda K_{ij}\eta_j = 0, \qquad \Phi_a(x, \eta, \eta') \equiv \Phi_{a\eta'_i}(x)\eta'_i + \Phi_{a\eta_i}(x)\eta_i = 0 \tag{2.4}$$

are satisfied; furthermore, there exist constants d_γ such that

$$\begin{aligned} & L_{i1}(\eta, \mu, d, \lambda) \\ & \qquad \equiv H\eta_{i1}[\eta] + d_\gamma\Psi_{\gamma\eta_{i1}} - \lambda G\eta_{i1}[\eta] - \Omega_{\eta'_i}(x, \eta, \eta', \mu) \,\big|^{x=x_1} = 0, \\ & L_{i2}(\eta, \mu, d, \lambda) \\ & \qquad \equiv H\eta_{i2}[\eta] + d_\gamma\Psi_{\gamma\eta_{i2}} - \lambda G\eta_{i2}[\eta] + \Omega_{\eta'_i}(x, \eta, \eta', \mu) \,\big|^{x=x_2} = 0, \\ & \Psi_\gamma[\eta(x_1), \eta(x_2)] \equiv \Psi_{\gamma\eta_{i1}}\eta_i(x_1) + \Psi_{\gamma\eta_{i2}}\eta_i(x_2) = 0. \end{aligned} \tag{2.5}$$

The boundary value problem (2.4), (2.5) is the one considered in this paper. A constant λ will be said to be a characteristic number if for this value there exist functions $\eta_i(x)$ of class C'' with multipliers $\mu_a(x)$ of class C' such that the set η_i, μ_a does not vanish identically on x_1x_2, the set satisfies (2.4) on this interval, and is such that there exist constants d_γ with which the end-values of the set satisfy (2.5).

(H3) The only solution η_i, μ_a of the system (2.4), (2.5) for which $\eta_i \equiv 0$ on x_1x_2 is the identically vanishing solution $\eta_i \equiv 0 \equiv \mu_a$ on x_1x_2. This hypothesis is a condition of normality on the interval x_1x_2 with respect to the differential equations (1.3) and the conditions (2.5), independent of the quantities $\| K_{ij} \|$ and $G[\eta(x_1), \eta(x_2)]$.†

† For a discussion of normality conditions for the problem of Lagrange, see G. A. Bliss, *American Journal of Mathematics*, Vol. 52 (1930), pp. 673-744; in particular, pp. 687-695. This paper will be referred to as B. I.

(H4) $I[\eta] > 0$ for every admissible arc η which is not identically zero on x_1x_2.

With respect to the boundary value problem (2.4), (2.5), condition (H4) is not as restrictive as it may first seem to be. If there is a constant λ_0 such that the expression

$$I[\eta] + \lambda_0\{2G[\eta(x_1), \eta(x_2)] + \int_{x_1}^{x_2} \eta_i K_{ij} \eta_j \, dx\}$$

is positive for all non-identically vanishing admissible arcs, then $I[\eta]$ may be replaced by this expression. The modified boundary value problem is equivalent to (2.4), (2.5), and may be reduced to it by a linear change of parameter.

3. Properties of the boundary value problem (2.4), (2.5). In this section will be given some fundamental properties of the differential system (2.4), (2.5). For brevity $H[\eta]$ will be written for $H[\eta(x_1), \eta(x_2)]$ and $H[u; v]$ will be used to denote $H\eta_{i1}[u]v_i(x_1) + H\eta_{i2}[u]v_i(x_2)$; corresponding notations will be introduced for $G[\eta(x_1), \eta(x_2)]$. We have the fundamental identities

$$\begin{aligned} H[u;\, v] &\equiv H[v;\, u], \qquad & H[u;\, u] &\equiv 2H[u], \\ G[u;\, v] &\equiv G[v;\, u], \qquad & G[u;\, u] &\equiv 2G[u]. \end{aligned} \tag{3.1}$$

Similarly, for the quadratic form $\Omega(x, \eta, \eta', \mu)$ we have:

$$2\Omega(x, \eta, \eta', \mu) \equiv \eta'_i \Omega_{\eta'_i} + \eta_i \Omega_{\eta_i} + \mu_a \Omega_{\mu_a}, \tag{3.2}$$

$$u'_i \Omega_{v'_i} + u_i \Omega_{v_i} + \rho_a \Omega_{\sigma_a} \equiv v'_i \Omega_{u'_i} + v_i \Omega_{u_i} + \sigma_a \Omega_{\rho_a}, \tag{3.3}$$

where the derivatives of Ω have the arguments (η, η', μ), (u, u', ρ) or (v, v', σ) as indicated by their subscripts.

It is also to be noted that if u_i, ρ_a is an arbitrary set of functions and d_γ and λ are arbitrary constants, then for every terminally admissible arc v we have

$$\begin{aligned} v_i(x_1) L_{i1}(u, \rho, d, \lambda) &+ v_i(x_2) L_{i2}(u, \rho, d, \lambda) \\ &\equiv H[u;\, v] - \lambda G[u;\, v] + v_i(x) \Omega_{u'_i}(x, u, u', \rho) \Big|_1^2 . \end{aligned} \tag{3.4}$$

In particular, if u_i, ρ_a is a set which satisfies (2.5) with constants d_γ and λ, then for every terminally admissible arc v we have

$$0 = H[u;\, v] - \lambda G[u;\, v] + v_i(x) \Omega_{u'_i}(x, u, u', \rho) \Big|_1^2 . \tag{3.4'}$$

If u and v are both differentially admissible arcs of class C'' and ρ_a is an arbitrary set of functions of class C', it follows from an integration by parts that

$$(3.5) \quad \int_{x_1}^{x_2} v_i J_i(u,\rho)\,dx = v_i \Omega_{u'_i}(x,u,u',\rho)\,\Big|_1^2 - \int_{x_1}^{x_2} [v'_i \Omega_{u'_i} + v_i \Omega_{u_i}]\,dx.$$

It follows by the use of (3.2), (3.4′) and (3.5) that

LEMMA 3.1. *If η_i, μ_a is a solution of the boundary value problem* (2.4), (2.5) *corresponding to a value λ, then the expression*

$$(3.6) \qquad I[\eta \mid \lambda] \equiv I[\eta] - \lambda\{2G[\eta] + \int_{x_1}^{x_2} \eta_i K_{ij} \eta_j\, dx\}$$

has the value zero.

COROLLARY. *The value $\lambda = 0$ is not a characteristic number of* (2.4), (2.5).

LEMMA 3.2. *If u_i, ρ_a and v_i, σ_a are solutions of* (2.4), (2.5) *corresponding to distinct characteristic numbers λ and λ^*, then*

$$G[u;\, v] + \int_{x_1}^{x_2} u_i K_{ij} v_j\, dx = 0.$$

This lemma follows directly from the relation

$$(\lambda - \lambda^*) \int_{x_1}^{x_2} u_i K_{ij} v_j dx = \int_{x_1}^{x_2} [u_i J_i(v,\sigma) - v_i J_i(u,\rho)]dx$$

and the relations (3.3), (3.4′) and (3.5).

LEMMA 3.3. *The boundary value problem* (2.4), (2.5) *has only real characteristic numbers.*

For suppose that η_i, μ_a is a solution of (2.4), (2.5) corresponding to a complex value λ. Then $\bar{\eta}_i$, $\bar{\mu}_a$, the conjugate imaginaries of the set η_i, μ_a, furnish a solution of this system corresponding to $\bar{\lambda}$, the conjugate imaginary of λ. By Lemma 3.2,

$$(3.7) \qquad G[\bar{\eta};\, \eta] + \int_{x_1}^{x_2} \bar{\eta}_i K_{ij} \eta_j\, dx = 0.$$

Furthermore,

$$\begin{aligned} 0 &= \int_{x_1}^{x_2} \bar{\eta}_i [J_i(\eta,\mu) + \lambda K_{ij}\eta_j]\, dx \\ (3.8) \qquad &= -H[\bar{\eta};\, \eta] - \int_{x_1}^{x_2} [\bar{\eta}'_i \Omega_{\eta'_i} + \bar{\eta}_i \Omega_{\eta_i}]\, dx, \end{aligned}$$

in view of (3.4′), (3.5) and (3.7). If u and v are the real and pure imaginary parts of η, it is readily shown that the right member of (3.8) reduces to $-I[u] - I[v]$, and therefore by (H4) we have that $u \equiv 0 \equiv v$, and hence $\eta \equiv 0$ on x_1x_2. But from (H3) it would then follow that the multipliers μ_a are all zero on x_1x_2, and therefore this value of λ is not a

characteristic number. Hence (2.4), (2.5) has only real characteristic numbers and the corresponding characteristic solutions may be chosen real.

4. The canonical form of the differential system (2.4), (2.5). In the equations (2.4) one may introduce the canonical variables

$$\zeta_i = \Omega_{\eta'_i}(x, \eta, \eta', \mu) \qquad (i = 1, 2, \cdots, n). \tag{4.1}$$

Since the matrix (2.1) is non-singular on x_1x_2, the system (1.3), (4.1) of $m + n$ equations has solutions

$$\eta'_i = \chi_i(x, \eta, \zeta); \qquad \mu_\alpha = M_\alpha(x, \eta, \zeta)$$

which are linear in η_i, ζ_i. Then the equations (2.4) are equivalent to

$$\eta'_i = \chi_i(x, \eta, \zeta), \quad \zeta'_i = \Omega_{\eta_i}[x, \eta, \chi(x, \eta, \zeta), M(x, \eta, \zeta)] - \lambda K_{ij}\eta_j. \tag{2.4$'$}$$

If $a_{\theta i}, b_{\theta i}$ $(\theta = 1, 2, \cdots, 2n - p)$ are linearly independent solutions of

$$\Psi_{\gamma\eta_{i1}} a_{\theta i} + \Psi_{\gamma\eta_{i2}} b_{\theta i} = 0 \qquad (\gamma = 1, 2, \cdots, p),$$

then (2.5) is equivalent to the $2n$ linearly independent relations

$$\begin{aligned} a_{\theta i}\{H_{\eta_{i1}} - \lambda G_{\eta_{i1}} - \zeta_i(x_1)\} + b_{\theta i}\{H_{\eta_{i2}} - \lambda G_{\eta_{i2}} + \zeta_i(x_2)\} &= 0 \\ \Psi_\gamma[\eta(x_1), \eta(x_2)] &= 0 \end{aligned} \tag{2.5$'$}$$

in the end-values of the variables η_i, ζ_i. The boundary value problem (2.4), (2.5) is then equivalent to the boundary value problem (2.4′), (2.5′) in the $2n$ variables η_i, ζ_i.

There will now be given some properties of the differential system (2.4′), (2.5′) which will be used in the later sections. One might proceed at once to these sections and then, whenever the results of this section are used, refer back to the present discussion.

To obtain compactness of notation we note that system (2.4′), (2.5′) is of the form

$$y'_\pi = (A_{\pi\tau} + \lambda B_{\pi\tau})y_\tau, \quad s_\pi(\lambda : y) \equiv M_{\pi\tau}(\lambda)y_\tau(x_1) + N_{\pi\tau}(\lambda)y_\tau(x_2) = 0, \quad (\pi, \tau = 1, 2, \cdots, 2n), \tag{4.2}$$

where $y_i = \eta_i$, $y_{n+i} = \zeta_i$ $(i = 1, 2, \cdots, n)$. The functions $A_{\pi\tau}$ and $B_{\pi\tau}$ are continuous on x_1x_2, while $M_{\pi\tau}$ and $N_{\pi\tau}$ are linear in λ and the matrix $\| M_{\pi\tau}(\lambda), N_{\pi\tau}(\lambda) \|$ has rank $2n$ for each value of λ. The general system (4.2) is of the type discussed by Bliss.† Let $Y(x, \lambda) \equiv \| Y_{\pi\tau}(x, \lambda) \|$ be

† G. A. Bliss, *Transactions of the American Mathematical Society*, Vol. 28 (1926), pp. 561-584. This paper will be referred to as B. II. In the system treated by Bliss $M_{\pi\tau}$ and $N_{\pi\tau}$ are supposed to be independent of λ, but the properties of the above system (4.2) that we state here may be proven by the same methods that Bliss uses.

a matrix whose columns are $2n$ linearly independent solutions of (4.2) and which reduces to the unit matrix for $x = x_1$. System (4.2) has exactly r linearly independent solutions for a given value λ if and only if the matrix $\| s_\pi[\lambda : Y_\tau(x, \lambda)] \|$ has rank $2n - r$. Now $| s_\pi[\lambda : Y_\tau(x, \lambda)] |$ is a permanently convergent power series in λ and its zeros are the characteristic numbers of (4.2). For the particular system (2.4′), (2.5′) we have that $\lambda = 0$ is not a characteristic number and therefore the corresponding determinant $| s_\pi[\lambda : Y_\tau(x, \lambda)] |$ is not identically zero. Hence for this system we have

LEMMA 4.1. *The totality of characteristic numbers of* (2.4′), (2.5′) *is denumerable and has no finite accumulation point.*

The system adjoint to (4.2) is

$$(4.3)\quad z'_\pi = -(A_{\tau\pi} + \lambda B_{\tau\pi})z_\tau, \quad t_\pi(\lambda : z) \equiv P_{\tau\pi}(\lambda) z_\tau(x_1) + Q_{\tau\pi}(\lambda) z_\tau(x_2) = 0,$$

where $P_{\tau\pi}$ and $Q_{\tau\pi}$ are such that the matrix of coefficients in the boundary conditions of (4.3) is of rank $2n$ and $M_{\pi\kappa}(\lambda) P_{\kappa\tau}(\lambda) - N_{\pi\kappa}(\lambda) Q_{\kappa\tau}(\lambda) = 0$ $(\pi, \tau = 1, 2, \cdots, 2n)$. For a given value of λ the number of linearly independent solutions is the same for (4.2) and (4.3).

If we define

$$s^*_\pi(\lambda : y) = M^*_{\pi\tau}(\lambda) y_\tau(x_1) + N^*_{\pi\tau}(\lambda) y_\tau(x_2)$$
$$t^*_\pi(\lambda : z) = P^*_{\tau\pi}(\lambda) z_\tau(x_1) + Q^*_{\tau\pi}(\lambda) z_\tau(x_2),$$

where $M^*_{\pi\tau}$, $N^*_{\pi\tau}$, $P^*_{\pi\tau}$, and $Q^*_{\pi\tau}$ are such that for each value of λ the matrices

$$\left\| \begin{matrix} M_{\pi\tau} & N_{\pi\tau} \\ M^*_{\pi\tau} & N^*_{\pi\tau} \end{matrix} \right\| \qquad \left\| \begin{matrix} -P^*_{\pi\tau} & -P_{\pi\tau} \\ Q^*_{\pi\tau} & Q_{\pi\tau} \end{matrix} \right\|$$

are reciprocals, then, as shown by Bliss, we have the identity

$$s_\pi(\lambda : y) t^*_\pi(\lambda : z) + s^*_\pi(\lambda : y) t_\pi(\lambda : z) = y_\pi(x) z_\pi(x) \Big|_{x=x_1}^{x=x_2}.$$

If $| s_\pi[\lambda : Y_\tau(x, \lambda)] | \neq 0$ for a given λ, then for this value of λ the non-homogeneous system

$$(4.4)\quad y'_\pi = (A_{\pi\tau} + \lambda B_{\pi\tau}) y_\tau + g_\pi, \qquad s_\pi(\lambda : y) = h_\pi \quad (\pi = 1, 2, \cdots, 2n),$$

where the functions $g_\pi(x)$ are continuous and the h_π are constants, has a unique solution of class C'. If the functions $g_\pi(x)$ are merely of class D, then (4.4) has a unique solution which is continuous and whose derivatives are continuous except possibly at the discontinuities of the functions $g_\pi(x)$. If $\| s_\pi[\lambda : Y_\tau(x, \lambda)] \|$ has rank $2n - r$ for a given λ, then corresponding to this value of λ the system (4.4) has a solution if and only if the relation

$$\int_{x_1}^{x_2} z_\pi(x) g_\pi(x)\, dx = h_\pi t^*_\pi(\lambda : z) \tag{4.5}$$

is satisfied by every solution z of the adjoint system (4. 3) for this value of λ.† The most general solution of (4. 4) is then $y_\pi = y^*_\pi + c_1 Y_{\pi 1} + \cdots + c_r Y_{\pi r}$, where (y^*_π) is a particular solution and $(Y_{\pi 1}), \cdots, (Y_{\pi r})$ are r linearly independent solutions of (4. 2).

Bliss has called a system (4. 2) self-adjoint if the differential equations and the boundary conditions of its adjoint are equivalent to its own for all values of λ by means of a transformation $z_\pi = T_{\pi\tau}(x) y_\tau$, where the matrix $\| T_{\pi\tau} \|$ is non-singular and its elements are of class C' on x_1x_2. The system (2. 4′), (2. 5′) may be shown to be self-adjoint, and the matrix of transformation $\| T_{\pi\tau} \|$ is the constant matrix ‡

$$\left\| \begin{matrix} 0 & \delta_{ij} \\ -\delta_{ij} & 0 \end{matrix} \right\| . \tag{4.6}$$

Suppose that corresponding to a given λ there are exactly r linearly independent solutions $(Y_{\pi 1}), \cdots, (Y_{\pi r})$ of (4. 2). Then solutions $(Y_{\pi, r+1}), \cdots, (Y_{\pi, 2n})$ may be chosen such that $Y(x, \lambda) \equiv \| Y_{\pi\tau}(x, \lambda) \|$ is non-singular on x_1x_2. Let $Z(x, \lambda) \equiv \| Z_{\pi\tau}(x, \lambda) \|$ be the reciprocal of $Y(x, \lambda)$ on x_1x_2; then each row of $Z(x, \lambda)$ is a solution of (4. 3). Now integers $i_1, i_2, \cdots, i_{2n-r}$ exist such that the matrix $\| s_{i\beta}[\lambda : Y_{r+\xi}(x, \lambda)] \|$ $(\beta, \xi = 1, 2, \cdots, 2n - r)$ has a unique reciprocal, which we will denote by $\| s_{\beta\xi}^{-1}(\lambda) \|$ $(\beta, \xi = 1, 2, \cdots, 2n - r)$. Now define the matrix $D(\lambda) \equiv \| D_{\pi\tau}(\lambda) \|$ as follows: $D_{r+\beta, i_\xi}(\lambda) = s_{\beta\xi}^{-1}(\lambda)$, $D_{\pi\tau}(\lambda) = 0$ if $\pi \leqq r$ or $\tau \neq i_\xi$ $(\beta, \xi = 1, 2, \cdots, 2n - r)$. Let

$$G_{\pi\tau}(x, t, \lambda) = \tfrac{1}{2} Y_{\pi\kappa}(x, \lambda) \left[\frac{|x - t|}{x - t} \delta_{\kappa\nu} + D_{\kappa v}(\lambda) \Delta_{v\nu}(\lambda) \right] Z_{\nu\tau}(t, \lambda), \tag{4.7}$$

where $Y_{\pi\tau}(x, \lambda)$, $Z_{\pi\tau}(x, \lambda)$ and $D_{\pi\tau}(\lambda)$ are defined as above and $\Delta_{\pi\tau}(\lambda) = M_{\pi\kappa}(\lambda) Y_{\kappa\tau}(x_1, \lambda) - N_{\pi\kappa}(\lambda) Y_{\kappa\tau}(x_2, \lambda)$. The subscripts ν and v which occur in (4. 7) are supposed to have also the range $1, 2, \cdots, 2n$. If $g_\pi(x)$ and h_π are such that the system (4. 4) has a solution for a given value of λ, then a particular solution of that system is given by

$$y_\pi(x, \lambda) = \int_{x_1}^{x_2} G_{\pi\tau}(x, t, \lambda) g_\tau(t)\, dt + Y_{\pi\kappa}(x, \lambda) D_{\kappa\tau}(\lambda) h_\tau . \S \tag{4.8}$$

---

† For the case $h_\pi = 0$ $(\pi = 1, 2, \cdots, 2n)$ this result is proven by Bliss. See B. II, p. 567. The same method of proof applies to the more general case to give the stated result.

‡ See Cope, *loc. cit.*

§ If $r = 0$, then $\| G_{\pi\tau}(x, t, \lambda) \|$ is the ordinary Green's matrix and the result is

The matrix $\| G_{\pi\tau}(x, t, \lambda) \|$ for (2.4′), (2.5′) we will denote by

$$(4.9) \qquad \left\| \begin{matrix} G^1_{ij}(x, t, \lambda) & G^2_{ij}(x, t, \lambda) \\ G^3_{ij}(x, t, \lambda) & G^4_{ij}(x, t, \lambda) \end{matrix} \right\|,$$

where $\| G^1_{ij} \|$, $\| G^2_{ij} \|$, $\| G^3_{ij} \|$ and $\| G^4_{ij} \|$ are n-rowed square matrices. If we apply the above result to the non-homogeneous system

$$(4.10) \quad \begin{gathered} \eta'_i = \chi_i(x, \eta, \zeta), \ \ \zeta'_i = \Omega_{\eta_i}[x, \eta, \chi(x, \eta, \zeta), M(x, \eta, \zeta)] - \lambda K_{ij}\eta_j + k_i, \\ a_{\theta i}\{H_{\eta_{i1}} - \lambda G_{\eta_{i1}} - \zeta_i(x_1)\} + b_{\theta i}\{H_{\eta_{i2}} - \lambda G_{\eta_{i2}} + \zeta_i(x_2)\} = h_\theta, \\ \Psi_\gamma[\eta(x_1), \eta(x_2)] = 0 \end{gathered}$$

corresponding to the homogeneous system (2.4′), (2.5′), we have that if this system has a solution for a given value of λ, then a particular solution is given by

$$(4.11) \quad \begin{aligned} \eta_i(x, \lambda) &= \int_{x_1}^{x_2} G^2_{ij}(x, t, \lambda) k_j(t)\, dt + Y_{i\kappa}(x, \lambda) D_{\kappa\theta}(\lambda) h_\theta, \\ \zeta_i(x, \lambda) &= \int_{x_1}^{x_2} G^4_{ij}(x, t, \lambda) k_j(t)\, dt + Y_{n+i,\kappa}(x, \lambda) D_{\kappa\theta}(\lambda) h_\theta. \end{aligned}$$

For the differential system (2.4′), (2.5′) we have in view of the corollary to Lemma 3.1 that $\| G_{\pi\tau}(x, t, 0) \|$ is an ordinary Green's matrix. Since this system is self-adjoint and the matrix $\| T_{\pi\tau} \|$ is given by (4.6), we have †

$$(4.12) \quad \begin{gathered} G^2_{ij}(x, t, 0) = G^2_{ji}(t, x, 0); \ \ G^3_{ij}(x, t, 0) = G^3_{ji}(t, x, 0); \\ G^1_{ij}(x, t, 0) = - G^4_{ji}(t, x, 0). \end{gathered}$$

From the form (4.11) for a solution of (4.10) if that system is compatible, we obtain at once that the differential system (2.4′), (2.5′) is equivalent to the following system of integral equations

$$(4.13) \quad \begin{aligned} \eta_i(x) &= -\lambda \int_{x_1}^{x_2} G^2_{ik}(x, t, 0) K_{kj}(t) \eta_j(t)\, dt + \lambda Y_{i\kappa}(x, 0) D_{\kappa\theta}(0) h_\theta^0, \\ \zeta_i(x) &= -\lambda \int_{x_1}^{x_2} G^4_{ik}(x, t, 0) K_{kj}(t) \eta_j(t)\, dt + \lambda Y_{n+i,\kappa}(x, 0) D_{\kappa\theta}(0) h_\theta^0, \end{aligned}$$

where $h_\theta^0 = a_{\theta i} G_{\eta_{i1}} + b_{\theta i} G_{\eta_{i2}}$ $(\theta = 1, 2, \cdots, 2n - p)$.

Finally, we state a lemma which will be of use in § 7,

---

proven by Bliss. See B. II. In this case the solution is unique. If $r > 0$, then $\| G_{\pi\tau}(x, t, \lambda) \|$ is a generalized Green's matrix and this result may be proven in the same manner as given in a recent paper by the author for the special case $h_\pi = 0$. See W. T. Reid, *American Journal of Mathematics*, Vol. 53 (1931), pp. 443-459; in particular, pp. 447, 448.

† See Bliss, B. II, p. 580. The relations (4.12) hold true for any value of λ which is not a characteristic number of (2.4′), (2.5′), but we use the relation only for the special case $\lambda = 0$.

LEMMA 4.2. *If the functions $k_j(x)$ are of class D on x_1x_2, and*

$$\eta_i(x) = \int_{x_1}^{x_2} G^2{}_{ij}(x, t, 0) k_j(t)\, dt, \tag{4.14}$$

then $\eta \equiv [\eta_i(x)]$ is an admissible arc.

This lemma follows immediately since η_i, ζ_i, where η_i is given by (4.14) and $\zeta_i(x) = \int_{x_1}^{x_2} G^4{}_{ij}(x, t, 0) k_j(t)\, dt$, is a solution of (4.10) for $\lambda = 0$, $h_\theta = 0$ $(\theta = 1, 2, \cdots, 2n - p)$.

5. Existence of characteristic numbers for (2.4), (2.5). There will now be defined a sequence of classes of admissible arcs, in which we shall consider the problem of minimizing the expression $I[\eta]$. The class S_1 is defined as consisting of the totality of admissible arcs η which satisfy the relation

$$2G[\eta] + \int_{x_1}^{x_2} \eta_i K_{ij} \eta_j\, dx = 1. \tag{5.1}$$

The sequence of classes is defined by induction as follows: Suppose classes $S_1, S_2, \cdots, S_{s-1}$ $(s \geqq 2)$ have been defined and are not empty, and for $\lambda = \lambda_t$ $(t = 1, 2, \cdots, s-1)$, where λ_t is the greatest lower bound of $I[\eta]$ in the class S_t, there are $r_t (0 < r_t \leqq 2n)$ linearly independent solutions of (2.4), (2.5). Let η_i^β, μ_α^β $(\beta = 1, 2, \cdots, r_1 + r_2 + \cdots + r_{s-1})$ denote these characteristic solutions of (2.4), (2.5), and let λ_β denote the value of λ corresponding to the solution η_i^β, μ_α^β. The class S_s is then defined as the totality of arcs η of class S_{s-1} which satisfy the relations

$$G[\eta^\beta; \eta] + \int_{x_1}^{x_2} \eta_i^\beta K_{ij} \eta_j\, dx = 0 \quad (\beta = 1, 2, \cdots, r_1 + \cdots + r_{s-1}). \tag{5.2}$$

There will be proven in this section the following theorem:

THEOREM 5.1. *If the class S_s is not empty and λ_s is the greatest lower bound of $I[\eta]$ in this class, then $\lambda = \lambda_s$ is a characteristic number of* (2.4), (2.5) *and $\lambda_s > \lambda_{s-1}$.*

In particular, the proof of this theorem applies to the case $s = 1$, when λ_0 is set equal to zero. By the corollary to Lemma 3.1 we know that $\lambda = 0$ is not a characteristic number of (2.4), (2.5), and we therefore obtain from Theorem 5.1 a complete induction proof for the existence of successive characteristic numbers for our boundary value problem.

Theorem 5.1 will be established by showing that if we assume it is false we are led to a contradiction. Clearly $\lambda_s \geqq \lambda_{s-1}$. If the theorem is false,

then either $\lambda_s > \lambda_{s-1}$ and is not a characteristic number of (2.4), (2.5), or $\lambda_s = \lambda_{s-1}$. Now let

$$(5.3) \qquad u_\nu \equiv (u_{i\nu}) \qquad (\nu = 1, 2, \cdots)$$

be a sequence of arcs of class S_s such that $\lim_{\nu\to\infty} I[u_\nu] = \lambda_s$. On the assumption that S_s is not empty such a sequence clearly exists. The following auxiliary lemma will now be proven.

LEMMA 5.1. *If Theorem 5.1 is false, then the non-homogeneous system*

$$(5.4) \qquad J_i(\eta, \mu) + \lambda_s K_{ij}\eta_j = K_{ij}u_{j\nu}, \qquad \Phi_a(x, \eta, \eta') = 0;$$

$$(5.5) \qquad L_{i1}(\eta, \mu, d, \lambda_s) + G\eta_{i1}[u_\nu] = 0 = L_{i2}(\eta, \mu, d, \lambda_s) + G\eta_{i2}[u_\nu] = \Psi_\gamma[\eta(x_1), (x_2)]$$

where u_ν is defined as above, has a solution $v_{i\nu}$, $\sigma_{a\nu}$ such that the arc v_ν satisfies the relations (5.2), *and furthermore,*

$$(5.6) \qquad |\, v_{i\nu}(x)\,| \leqq l_1 \sum_{i=1}^{n} \{ \int_{x_1}^{x_2} u^2_{i\nu}dx + u^2_{i\nu}(x_1) + u^2_{i\nu}(x_2)\} + l_2 \qquad (\nu = 1, 2, \cdots)$$

on x_1x_2, where l_1 and l_2 are constants independent of x and ν.

By the introduction of the canonical variables (4.1) the system (5.4), (5.5) is reduced to the system (4.10), with $\lambda = \lambda_s$, $k_i = K_{ij}u_{j\nu}$ and $h_\theta = -\{G\eta_{i1}[u_\nu]a_{\theta i} + G\eta_{i2}[u_\nu]b_{\theta i}\}$. If $\lambda_s > \lambda_{s-1}$ and is not a characteristic number of (2.4), (2.5), then there exists a unique solution of (5.4), (5.5); if $\lambda_s = \lambda_{s-1}$, since the arcs u_ν satisfy (5.2) it may be shown that the functions $k_i(x)$, h_θ defined above satisfy for the system (4.10) the condition (4.5) for the value $\lambda = \lambda_s$, and hence the system (4.10) is still compatible. It then follows that for $\lambda = \lambda_s$, $k_i = K_{ij}u_{j\nu}$, $h_\theta = -\{G\eta_{i1}[u_\nu]a_{\theta i} + G\eta_{i2}[u_\nu]b_{\theta i}\}$ the system (4.10) has a particular solution $\eta^0{}_{i\nu}$, $\zeta^0{}_{i\nu}$ given by the relations (4.11), and by the application of elementary inequalities it is seen that the functions $\eta^0{}_{i\nu}(x)$ satisfy an inequality of the form (5.6). Now if $\sigma^0{}_{a\nu} = M_a(x, \eta_\nu{}^0, \zeta_\nu{}^0)$, we have

$$(\lambda_s - \lambda_\beta) \int_{x_1}^{x_2} \eta^0{}_{i\nu}K_{ij}\eta_j{}^\beta dx = \int_{x_1}^{x_2} [\eta^0{}_{i\nu}J_i(\eta^\beta, \mu^\beta) - \eta_i{}^\beta J_i(\eta_\nu{}^0, \sigma_\nu{}^0) + \eta_i{}^\beta K_{ij}u_{j\nu}]dx,$$

and since the arc u_ν is of class S_s, it follows in view of (3.3), (3.4) and (3.5) that

$$(5.7) \qquad (\lambda_s - \lambda_\beta)\{ \int_{x_1}^{x_2} \eta^0{}_{i\nu}K_{ij}\eta_j{}^\beta dx + G[\eta_\nu{}^0;\ \eta^\beta]\} = 0 \qquad (\beta = 1, 2, \cdots, r_1 + \cdots + r_{s-1}).$$

Now set

$$
(5.8)\qquad
\begin{aligned}
\eta^*_{i\nu} &= \eta^0_{i\nu} - \sum_{\beta=1}^{R} \{ \int_{x_1}^{x_2} \eta^0_{i\nu} K_{ij} \eta_j^\beta dx + G[\eta_\nu^0;\, \eta^\beta] \} \eta_i^\beta, \\
\zeta^*_{i\nu} &= \zeta^0_{i\nu} - \sum_{\beta=1}^{R} \{ \int_{x_1}^{x_2} \eta^0_{i\nu} K_{ij} \eta_j^\beta dx + G[\eta_\nu^0;\, \eta^\beta] \} \zeta_i^\beta,
\end{aligned}
$$

where $R = r_1 + , \cdots , r_{s-1}$. If $\lambda_s > \lambda_{s-1}$, in view of (5.7) we have that $\eta^*_{i\nu}, \zeta^*_{i\nu} \equiv \eta^0_{i\nu}, \zeta^0_{i\nu}$. If $\lambda_s = \lambda_{s-1}$, then by (5.7) the coefficients of $\eta_i^\beta(x)$ and $\zeta_i^\beta(x)$ in (5.8) are zero unless $\lambda_\beta = \lambda_{s-1}$. In either case the set $\eta^*_{i\nu}, \zeta^*_{i\nu}$ is a solution of (4.10) for the values of λ, k_i and h_θ indicated above, and the functions $\eta^*_{i\nu}$ satisfy (5.2) with all the functions η_i^β $(\beta = 1, 2, \cdots, r_1 + \cdots + r_{s-1})$. Since the functions $\eta^0_{i\nu}(x)$ satisfy an inequality of the form (5.6) it is readily verified that the functions $\eta^*_{i\nu}$ satisfy a similar inequality. The functions $v_{i\nu} = \eta^*_{i\nu}$, together with the multipliers $\rho_{a\nu}$ and constants $d_{\gamma\nu}$ corresponding to the solution $\eta^*_{i\nu}, \zeta^*_{i\nu}$ of (4.10), give a solution of (5.4), (5.5) which satisfies the relations of Lemma 5.1.

Now let

$$
(5.9)\qquad w_{i\nu}(x) = u_{i\nu}(x) + c v_{i\nu}(x) \qquad (i = 1, 2, \cdots, n;\ \nu = 1, 2, \cdots),
$$

where c is a real constant. Then w_ν satisfies relations (5.2), and

$$
\begin{aligned}
I[w_\nu \mid \lambda_s] &= I[u_\nu \mid \lambda_s] + c^2 I[v_\nu \mid \lambda_s] + 2c\big(H[u_\nu;\, v_\nu] - \lambda_s\, G[u_\nu;\, v_\nu] \\
&\quad + \int_{x_1}^{x_2} \{u'_{i\nu}\Omega_{\eta'_i}(x, v_\nu, v'_\nu, \sigma_\nu) + u_{i\nu}\Omega_{\eta_i}(x, v_\nu, v'_\nu, \sigma_\nu) - \lambda_s u_{i\nu} K_{ij} v_{j\nu}\} dx\big).
\end{aligned}
$$

When the integrals which occur in the coefficients of c and c^2 are integrated by parts in the usual manner and use is made of the fact that $v_{i\nu}, \sigma_{a\nu}$ is a solution of (5.4), (5.5) and that the functions u_ν are of class S_s, one obtains

$$
\begin{aligned}
(5.10)\qquad I[w_\nu \mid \lambda_s] &= I[u_\nu \mid \lambda_s] - 2c \\
&\quad - c^2\{G[u_\nu;\, v_\nu] + \int_{x_1}^{x_2} u_{i\nu} K_{ij} v_{j\nu} dx\} \qquad (\nu = 1, 2, \cdots).
\end{aligned}
$$

From the above relation we obtain the following lemma, which is of use in proving the general Theorem 5.1.

LEMMA 5.2. *There exists a positive constant l such that for every admissible arc η we have*

$$
I[\eta] \geqq l\{ \int_{x_1}^{x_2} \eta_i \eta_i dx + \eta_i(x_1)\eta_i(x_1) + \eta_i(x_2)\eta_i(x_2) \}.
$$

The proof of this lemma depends upon the above relations for the special case in which $\| K_{ij} \|$ is the unit matrix and $2G[\eta(x_1), \eta(x_2)]$

$=\eta_i(x_1)\eta_i(x_1)+\eta_i(x_2)\eta_i(x_2)$, while $s=1$. Let $S_1{}^0$ denote the above defined class S_1 corresponding to these particular values of $\| K_{ij} \|$ and $G[\eta]$. Lemma 5.2 then states that $\lambda_1{}^0>0$, where $\lambda_1{}^0$ is the greatest lower bound of $I[\eta]$ in the class $S_1{}^0$. Suppose that the lemma were not true. Since $\lambda_1{}^0 \geqq 0$ we would have $\lambda_\nu{}^0=0$, and there would exist a sequence of arcs $u_\nu{}^0$ $(\nu=1,2,\cdots)$ of class $S_1{}^0$ such that $\lim_{\nu\to\infty} I[u_\nu{}^0]=0$. Since (2.4), (2.5) is not compatible for $\lambda=0$, we have by Lemma 5.1 that the solution $v^0{}_{i\nu}, \sigma^0{}_{a\nu}$ of (5.4), (5.5) for $\lambda=0$ and corresponding to $u_\nu{}^0$ for these particular values of $\| K_{ij} \|$ and $G[\eta]$, satisfies a relation

$$\begin{aligned}|v^0{}_{i\nu}(x)| &\leqq l_3 \sum_{i=1}^{n} \{ \int_{x_1}^{x_2} (u^0{}_{i\nu})^2\,dx + [u^0{}_{i\nu}(x_1)]^2 + [u^0{}_{i\nu}(x_2)]^2\} + l_4,\\ &\leqq l_3 + l_4,\end{aligned}$$

where l_3 and l_4 are constants independent of x and ν. For $w^0{}_{i\nu}(x)=u^0{}_{i\nu}(x)+cv^0{}_{i\nu}(x)$ $(\nu=1,2,\cdots)$ we have from (5.10) that

$$\begin{aligned}I[w_\nu{}^0] &= I[u_\nu{}^0]-2c\\ &\quad - c^2 \sum_{i=1}^{n} \{u^0{}_{i\nu}(x_1)v^0{}_{i\nu}(x_1)+u^0{}_{i\nu}(x_2)v^0{}_{i\nu}(x_2)+\int_{x_1}^{x_2} u^0{}_{i\nu}v^0{}_{i\nu}dx\}\end{aligned}$$

Now the coefficient of c^2 in this expression is in absolute value not greater than

$$\begin{aligned}\tfrac{1}{2} \sum_{i=1}^{n} \{&[u^0{}_{i\nu}(x_1)]^2+[u^0{}_{i\nu}(x_2)]^2\\ &+[v^0{}_{i\nu}(x_1)]^2+[v^0{}_{i\nu}(x_2)]^2+\int_{x_1}^{x_2}[(u^0{}_{i\nu})^2+(v^0{}_{i\nu})^2]\,dx\}\end{aligned}$$

and since $u_\nu{}^0$ is of class $S_1{}^0$ and $v_\nu{}^0$ satisfies the above inequality we have that there is a positive constant l_5 such that the coefficient of c^2 in the expression for $I[w_\nu{}^0]$ is in absolute value less than l_5. Since $\lim_{\nu\to\infty} I[u_\nu{}^0]=0$, if we choose c such that $c^2 l_5-2c<0$ it then follows that for ν sufficiently large we have $I[w_\nu{}^0]<0$, which is impossible in view of (H4). Hence $\lambda_1{}^0>0$ and for $0<l\leqq\lambda_1{}^0$ we have the inequality of Lemma 5.2 satisfied.

Let us now apply the results of these lemmas to prove the general Theorem 5.1. On the assumption that the theorem is false we have corresponding to each arc u_ν of (5.3) a solution $v_{i\nu}, \sigma_{a\nu}$ of (5.4), (5.5) satisfying the relations of Lemma 5.1, and finally for w_ν defined by (5.9) we have relation (5.10). Now clearly there exists a positive constant l_6 such that

$$\begin{aligned}|G[u_\nu;v_\nu]+\int_{x_1}^{x_2} u_{i\nu}K_{ij}v_{j\nu}dx| &\leqq l_6 \sum_{i=1}^{n} \{u^2{}_{i\nu}(x_1)+u^2{}_{i\nu}(x_2)\\ &\quad +v^2{}_{i\nu}(x_1)+v^2{}_{i\nu}(x_2)+\int_{x_1}^{x_2}[u^2{}_{i\nu}(x)+v^2{}_{i\nu}(x)]\,dx\}.\end{aligned}$$

Since $\lim_{\nu\to\infty} I[u_\nu \mid \lambda_s] = \lim_{\nu\to\infty} \{I[u_\nu] - \lambda_s\} = 0$, it follows in view of 5.6 and Lemma 5.2 that there exists a positive constant l_7 such that the coefficient of c^2 in (5.10) is in absolute value less than l_7. Then, as in the proof of Lemma 5.2, we have that if c is chosen such that $c^2 l_7 - 2c < 0$ then for sufficiently large values of ν it follows that $I[w_\nu \mid \lambda_s] < 0$, which is impossible since λ_s was chosen as the greatest lower bound of $I[\eta]$ in the class S_s. Hence $\lambda = \lambda_s$ is a characteristic number of (2.4), (2.5) and $\lambda_s > \lambda_{s-1}$.

We shall now define the class S_{-1} as consisting of the totality of admissible arcs η such that

$$2G[\eta] + \int_{x_1}^{x_2} \eta_i K_{ij} \eta_j \, dx = -1, \tag{5.11}$$

and a sequence of classes $S_{-1}, S_{-2}, \cdots$ may be defined by induction in a manner analogous to that used in defining the sequence $S_1, S_2, \cdots$. If the matrix $\| K_{ij} \|$ is replaced by $\| -K_{ij} \|$ and $G[\eta]$ by $-G[\eta]$, then the classes $S_{-1}, S_{-2}, \cdots$ of the original problem correspond to the classes $S_1, S_2, \cdots$ for the modified problem. The following result then follows as a corollary to Theorem 5.1.

THEOREM 5.2. *If the class S_{-s} is not empty and $-\lambda_{-s}$ is the greatest lower bound of $I[\eta]$ in this class, then $\lambda = \lambda_{-s}$ is a negative characteristic number of* (2.4), (2.5) *and $\lambda_{-s} < \lambda_{-(s-1)}$.*

We have also the following property of the characteristic solutions of (2.4), (2.5).

LEMMA 5.3. *If g is an admissible arc and $G[\eta; g] + \int_{x_1}^{x_2} \eta_i K_{ij} g_j \, dx = 0$ for every solution η_i, μ_a of* (2.4), (2.5), *then $2G[g] + \int_{x_1}^{x_2} g_i K_{ij} g_j \, dx = 0$.*

For if the lemma were not true there would exist an admissible arc g satisfying the condition of the lemma with every solution η_i, μ_a of (2.4), (2.5) and $2G[g] + \int_{x_1}^{x_2} g_i K_{ij} g_j \, dx - l_8 \neq 0$; for definiteness, suppose $l_8 > 0$. There would then exist an arc η of class S_1 which satisfies the condition of the lemma, and for which $I[\eta] = l_9 > 0$. It readily follows that there would be infinitely many positive characteristic numbers of (2.4), (2.5) not greater than l_9, which is impossible in view of Lemma 4.1. A similar contradiction is obtained if we suppose that the arc g is such that $l_8 < 0$. Hence the lemma is proven.

From the results of this section it follows that each characteristic number

of the boundary value problem (2.4), (2.5) is uniquely determined as the minimum value of $I[\eta]$ in a corresponding class of admissible arcs.

6. Sufficient conditions for the existence of infinitely many characteristic numbers for (2.4), (2.5). On the assumption that the classes of functions $S_1, S_2, \cdots, S_s$ defined in § 5 are not empty it has been proven that corresponding characteristic numbers $0 < \lambda_1 < \cdots < \lambda_s$ exist for the boundary value problem (2.4), (2.5). If there exists an infinite sequence of classes $S_1, S_2, \cdots$ which are not empty there will clearly exist infinitely many positive characteristic numbers for (2.4), (2.5). A corresponding statement is true concerning the existence of infinitely many negative characteristic numbers. In this section will be given sufficient conditions for the boundary value problem (2.4), (2.5) to have infinitely many characteristic numbers.

Suppose that the matrix $\| K_{ij} \|$ and the differential equations (1.3) satisfy the additional condition:

($\mathrm{H}5^+$) There is a sub-interval $x'_1 x'_2$ of $x_1 x_2$ such that if $\bar{x}_1$ and $\bar{x}_2$ are distinct points so that $x'_1 \leqq \bar{x}_1 < \bar{x}_2 \leqq x'_2$, then there exists a differentially admissible arc η which vanishes at $\bar{x}_1$ and $\bar{x}_2$ and such that $\int_{\bar{x}_1}^{\bar{x}_2} \eta_i K_{ij} \eta_j dx > 0$.

THEOREM 6.1. *If hypotheses* (H1), (H2), (H3), (H4) *and* ($\mathrm{H}5^+$) *are satisfied, then the boundary value problem* (2.4), (2.5) *has infinitely many positive characteristic numbers.*

It will first be shown that the class S_1 is not empty. If $x'_1 x'_2$ is a subinterval which satisfies the condition of ($\mathrm{H}5^+$), then there exists a differentially admissible arc $u \equiv (u_i)$ which vanishes at x'_1 and x'_2 and is such that $\int_{x'_1}^{x'_2} u_i K_{ij} u_j dx > 0$. If u is defined as equal to zero outside the subinterval $x'_1 x'_2$, then $2G[u] + \int_{x_1}^{x_2} u_i K_{ij} u_j dx > 0$ and therefore S_1 is not empty.

It will now be proven that if the classes $S_1, S_2, \cdots, S_{s-1}$ are not empty then the class S_s is also not empty. We shall use the notation of § 5. Let $R = r_1 + \cdots + r_{s-1}$ and divide $x'_1 x'_2$ into $R + 1$ consecutive intervals. By ($\mathrm{H}5^+$) there exists an admissible arc $u_q \equiv (u_{iq})$ which vanishes outside the q-th interval and such that $\int_{x_1}^{x_2} u_{iq} K_{ij} u_{jq} dx > 0$. Then clearly constants $\delta_1, \delta_2, \cdots, \delta_{R+1}$ which are not all zero may be chosen such that the admissible arc defined as $u_i = \sum_{q=1}^{R+1} u_{iq} \delta_q$ satisfies the R linear relations $\int_{x_1}^{x_2} \eta_i{}^\beta K_{ij} u_j dx = 0$

$(\beta = 1, 2, \cdots, R)$. Now $2G[u] + \int_{x_1}^{x_2} u_i K_{ij} u_j dx > 0$, and therefore the class S_s is not empty. Theorem 6.1 is therefore established by induction.

Let (H5⁻) denote the hypothesis obtained by replacing in (H5⁺) the relation "$\int_{\bar{x}_1}^{\bar{x}_2} \eta_i K_{ij} \eta_j dx > 0$" by "$\int_{\bar{x}_1}^{\bar{x}_2} \eta_i K_{ij} \eta_j dx < 0$". This is equivalent to replacing $\| K_{ij} \|$ by $\| -K_{ij} \|$. We then have

COROLLARY. *If the hypotheses* (H1), (H2), (H3), (H4) *and* (H5⁻) *are satisfied, then the boundary value problem* (2.4), (2.5) *has infinitely many negative characteristic numbers.*

Let (H2′) and (H3′) denote the following strengthened hypotheses:

(H2′) In addition to the matrix (2.1) being non-singular, the quadratic form

$$\omega_{\eta'_i \eta'_j}(x) \pi_i \pi_j \qquad x_1 \leqq x \leqq x_2$$

is non-negative for every set $(\pi_i) \neq (0)$ which satisfies the equations

$$\Phi_{\alpha \eta'_j} \pi_j = 0 \qquad (\alpha = 1, 2, \cdots, m).$$

(H3′) If $x'_1 x'_2$ is any sub-interval of $x_1 x_2$, then there do not exist functions $\mu_\alpha(x)$ not identically zero and constants c_i such that

$$\mu_\alpha(x) \Phi_{\alpha \eta'_i}(x) = \int_{x'_1}^{x} \mu_\alpha(t) \Phi_{\alpha \eta_i}(t) dt + c_i \quad x'_1 \leqq x \leqq x'_2 \quad (i = 1, 2, \cdots, n).$$

This assumption of normality on every sub-interval with respect to the differential equations (1.3) clearly implies (H3).

THEOREM 6.2. *If the system* (2.4), (2.5) *satisfies* (H1), (H2′), (H3′), *while the quadratic form* $\eta_i K_{ij}(x) \eta_j$ *is positive definite for each* x *on* $x_1 x_2$ *and* $G[\eta]$ *is non-negative, then for this boundary value problem there exist infinitely many positive characteristic numbers and only a finite number of negative characteristic numbers.*

If $x'_1 x'_2$ is any sub-interval of $x_1 x_2$, it follows from (H3′) that there are infinitely many differentially admissible arcs which vanish at x'_1 and x'_2 and are not identically zero on $x'_1 x'_2$.† Since $\eta_i K_{ij}(x) \eta_j$ is positive definite, condition (H5⁺) is clearly satisfied. If (H4) is satisfied then Theorem 6.1 tells us that the boundary value problem has infinitely many positive characteristic numbers, and, since in this case the class S_{-1} is seen to be empty, there could be no negative characteristic numbers. If (H4) is not satisfied, however, in view of (H2′) and the condition that $\eta_i K_{ij}(x) \eta_j$ is positive

† See B. I, p. 689.

definite and $G[\eta]$ is non-negative, it follows that there exists a positive constant λ_0 such that the relation

$$0 \leqq I[\eta] + \lambda_0 \int_{x_1}^{x_2} \eta_i K_{ij} \eta_j dx \leqq I[\eta] + \lambda_0 \{2G[\eta] + \int_{x_1}^{x_2} \eta_i K_{ij} \eta_j dx\} \dagger$$

is satisfied by every admissible arc η, and the equality holds only if $\eta_i \equiv 0$ on $x_1 x_2$. Theorem 6.2 then follows in view of the remark made at the end of § 2.

The denumerable set of characteristic solutions and characteristic numbers of the boundary value problem (2.4), (2.5) may be represented by the symbols $\eta_{is}(x)$, $\mu_{as}(x)$, λ_s $(s = 1, 2, \cdots)$. These solutions may also be chosen orthonormal in the sense that

$$(6.1) \qquad G[\eta_s;\ \eta_t] + \int_{x_1}^{x_2} \eta_{is} K_{ij} \eta_{jt} dx = \delta_{st} (\operatorname{sgn} \lambda_s) \qquad (s, t = 1, 2, \cdots),$$

where $\delta_{st} = 0$ if $s \neq t$, $\delta_{ss} = 1$; $\operatorname{sgn} \lambda_s = 1$ if $\lambda_s > 0$, $\operatorname{sgn} \lambda_s = -1$ if $\lambda_s < 0$.

7. Expansion theorems. In this section there will be considered some expansion theorems in terms of the characteristic solutions of the boundary value problem (2.4), (2.5) *for the special case in which the coefficients of* $G[\eta(x_1), \eta(x_2)]$ *are zero. It will be assumed throughout this section that the hypotheses* (H1), (H2), (H3) *and* (H4) *are satisfied and that* (2.4), (2.5) *has infinitely many characteristic numbers.* The denumerably infinite set of characteristic solutions and characteristic numbers will be represented by $\eta_{is}(x)$, $\mu_{as}(x)$, λ_s $(s = 1, 2, \cdots)$ and we shall suppose that these solutions are orthonormal in the sense defined at the close of the last section; that is, such that

$$(7.1) \qquad \int_{x_1}^{x_2} \eta_{is} K_{ij} \eta_{jt} dx = \delta_{st} (\operatorname{sgn} \lambda_s) \qquad (s, t = 1, 2, \cdots).$$

The canonical variables corresponding to the solution η_{is}, μ_{as}, and defined by (4.1), will be denoted by ζ_{is}.

We shall also make the following additional hypothesis:

(H6) If g is an arbitrary admissible arc and

$$(7.2) \qquad \int_{x_1}^{x_2} \eta_{is} K_{ij} g_j dx = 0 \qquad (s = 1, 2, \cdots),$$

then $K_{ij}(x) g_j(x) \equiv 0$ on $x_1 x_2$.

† See M, p. 533. The proof is given for $\| K_{ij} \|$ the identity matrix, but the same method applies to the case in which $\eta_i K_{ij} \eta_j$ is positive definite. The method used by Hickson (*loc. cit.*, p. 570) to establish the existence of such a value λ_0 for the special case of our general problem which he considered, may also be extended to obtain the above relation.

If the quadratic form $\eta_i K_{ij}\eta_j$ is positive definite on x_1x_2, then hypothesis (H6) is a consequence of the preceding hypotheses, and follows as a corollary to Lemma 5.3. We will give, however, another condition that will insure (H6).

LEMMA 7.1. *Hypothesis* (H6) *is satisfied by the characteristic solutions of* (2.4), (2.5) *whenever the following condition is satisfied by the differential system:*

(C) *If g is an arbitrary admissible arc and η_i, μ_a is the solution of the system*

$$
(7.3)\qquad \begin{array}{l} J_i(\eta,\mu)+K_{ij}g_j=0, \qquad \Phi_a(x,\eta,\eta')=0, \\ L_{i1}(\eta,\mu,d,0)=0=L_{i2}(\eta,\mu,d,0)=\Psi_\gamma[\eta(x_1),\eta(x_2)], \end{array}
$$

where L_{i1} and L_{i2} are defined by (2.5), *then $K_{ij}(x)\eta_j(x)\equiv 0$ on x_1x_2 only if $\eta_i\equiv 0\equiv\mu_a$* on x_1x_2.

For if the lemma were false there would exist an admissible arc g satisfying the conditions (7.2), and such that the functions $K_{ij}(x)g_j(x)$ are not all identically zero on x_1x_2. For such an arc g the non-homogeneous system (7.3) would have a unique solution u_i, ρ_a since $\lambda=0$ is not a characteristic number of (2.4), (2.5); furthermore, in view of condition (C), the functions $K_{ij}(x)u_j(x)$ are not all identically zero on x_1x_2. For such a solution u_i, ρ_a of (7.3) we would have in view of (3.3), (3.4′) and (3.5) that

$$
(7.4)\qquad \lambda_s\int_{x_1}^{x_2}u_iK_{ij}\eta_{js}dx-\int_{x_1}^{x_2}\eta_{is}K_{ij}g_jdx
$$
$$
=\int_{x_1}^{x_2}[\eta_{is}J_i(u,\rho)-u_iJ_i(\eta_s,\mu_s)]\,dx=0\qquad(s=1,2,\cdots),
$$

and since $\lambda=0$ is not a characteristic number of (2.4), (2.5) the arc u would also satisfy (7.2). If in the differential equations of (7.3) the arc g were replaced by the arc u, this system would have a unique solution v_i, σ_a and again, in view of (C), the functions $K_{ij}(x)v_j(x)$ are not all identically zero on x_1x_2; furthermore, the arc v would also satisfy the relations (7.2). Now set $w_i=v_i-u_i$ $(i=1,2,\cdots,n)$. The arc w would satisfy (7.2), and it would follow from Lemma 5.3 that

$$
0=\int_{x_1}^{x_2}[w_iK_{ij}w_j-u_iK_{ij}u_j-v_iK_{ij}v_j]\,dx
$$
$$
=-2\int_{x_1}^{x_2}v_iK_{ij}u_jdx
$$

$$= 2 \int_{x_1}^{x_2} v_i J_i(v, \sigma)\, dx$$
$$= -2I[v],$$

and therefore, in view of (H4), $v_i \equiv 0$, which is a contradiction. Hence (C) is a sufficient condition to insure (H6).

Condition (C) is related to the condition that Bliss imposes in his definition of a definitely self-adjoint boundary value problem.† If (2.4), (2.5) is definitely self-adjoint then condition (C) is satisfied, but the converse is readily seen to not be true. The expansion theorems that we state will be related to the solutions of a non-homogeneous system of the form (7.3).

For an arbitrary arc g, define

$$(7.5) \qquad c_s[g] = \operatorname{sgn} \lambda_s \int_{x_1}^{x_2} g_i K_{ij} \eta_{js}\, dx \qquad (s = 1, 2, \cdots).$$

THEOREM 7.1. *If g is an arbitrary admissible arc such that the series*

$$(7.6) \qquad \sum_{s=1}^{\infty} c_s[g]\eta_{is}(x), \qquad \sum_{s=1}^{\infty} c_s[g]\zeta_{is}(x)$$

converge absolutely and uniformly on x_1x_2, and u_i, ρ_a is the solution of (7.3) corresponding to this arc g, then the series

$$(7.7) \qquad \sum_{s=1}^{\infty} c_s[u]\eta_{is}(x), \qquad \sum_{s=1}^{\infty} c_s[u]\zeta_{is}(x)$$

converge absolutely and uniformly on x_1x_2 to the values $u_i(x)$, $\Omega_{\eta'_i}[x, u(x), u'(x), \rho(x)]$.

From (7.4) it follows that

$$\lambda_s c_s[u] = c_s[g] \qquad (s = 1, 2, \cdots),$$

and therefore the series (7.7) converge absolutely and uniformly on x_1x_2. Now the set $g_i - \sum_{s=1}^{\infty} c_s[g]\eta_{is}(x)$ satisfies the relations (7.2), and therefore $K_{ij}\{g_j - \sum_{s=1}^{\infty} c_s[g]\eta_{js}\} \equiv 0$ on x_1x_2. It then follows that the set $u_i - \sum_{s=1}^{\infty} c_s[u]\eta_{is}$, $\Omega_{\eta'_i}(x, u, u', \rho) - \sum_{s=1}^{\infty} c_s[u]\zeta_{is}$ satisfies the system (4.10) for $\lambda = 0$, $k_i = 0$, $h_\theta = 0$, and is therefore identically zero. It also follows that the series $\sum_{s=1}^{\infty} c_s[u]\eta'_{is}(x)$ and $\sum_{s=1}^{\infty} c_s[u]\mu_{as}(x)$ converge absolutely and uniformly on x_1x_2 to the functions $u'_i(x)$, $\rho_a(x)$.

† See B. II, p. 570.

To prove a further expansion theorem for our system we make use of the following lemmas:

LEMMA 7.2. *If g is an admissible arc, then*

$$\sum_{s=1}^{\infty} c_s^2[g] \mid \lambda_s \mid \leqq I[g]. \tag{7.8}$$

This lemma follows from (H4) and the fact that if N is any positive integer one may prove directly that

$$I\{g - \sum_{s=1}^{N} c_s[g]\eta_s\} = I[g] - \sum_{s=1}^{N} c_s^2[g] \mid \lambda_s \mid.$$

LEMMA 7.3. *If g is an admissible arc and u_i, ρ_a is the solution of* (7.3) *corresponding to g, then the series* (7.7) *converge absolutely and uniformly on x_1x_2.*

It follows from (4.13) that

$$\eta_{is}(x) = -\lambda_s \int_{x_1}^{x_2} G^2{}_{il}(x,t) K_{lj}(t) \eta_{js}(t)\, dt, \tag{7.9}$$

$$\zeta_{is}(x) = -\lambda_s \int_{x_1}^{x_2} G^4{}_{il}(x,t) K_{lj}(t) \eta_{js}(t)\, dt, \tag{7.10}$$

where we have written $G^2{}_{il}(x,t)$ and $G^4{}_{il}(x,t)$ in place of $G^2{}_{il}(x,t,0)$ and $G^4{}_{il}(x,t,0)$. We then have in view of (4.12) that

$$\begin{aligned} \eta_{is}(x) &= \lambda_s^2 \int_{x_1}^{x_2} G^2{}_{il}(x,t) K_{lk}(t) [\int_{x_1}^{x_2} G^2{}_{kr}(t,\xi) K_{rj}(\xi) \eta_{js}(\xi)\, d\xi]\, dt \\ &= \lambda_s^2 \int_{x_1}^{x_2} \eta_{js}(\xi) K_{jr}(\xi) [\int_{x_1}^{x_2} G^2{}_{rk}(\xi,t) K_{kl}(t) G^2{}_{li}(t,x)\, dt]\, d\xi. \end{aligned}$$

In these expressions the subscripts l and r have also the range $1, 2, \cdots, n$. We likewise obtain

$$\zeta_{is}(x) = \lambda_s^2 \int_{x_1}^{x_2} \eta_{js}(\xi) K_{jr}(\xi) [\int_{x_1}^{x_2} G^2{}_{rk}(\xi,t) K_{kl}(t) G^4{}_{il}(x,t)\, dt]\, d\xi.$$

From Lemmas 4.2 and 7.2 we then obtain that for each fixed value of x the series

$$\sum_{s=1}^{\infty} \left(\frac{\eta_{is}(x)}{\lambda_s^2}\right)^2 \mid \lambda_s \mid, \qquad \sum_{s=1}^{\infty} \left(\frac{\zeta_{is}(x)}{\lambda_s^2}\right)^2 \mid \lambda_s \mid \tag{7.11}$$

converge; furthermore, from the explicit form (4.7) of the Green's matrix it is seen that there exists a constant l_{10} such that the series are less than l_{10}, uniformly for x on x_1x_2. Now since $\lambda_s c_s[u] = c_s[g]$, we have for each pair of integers N_1, N $(N_1 < N)$ the following relation

$$\sum_{s=N_1}^{N} c_s[u]\eta_{is}(x) = \sum_{s=N_1}^{N} \lambda_s \frac{\eta_{is}(x)}{\lambda_s^2} c_s[g]$$

and therefore

$$\left\{ \sum_{s=N_1}^{N} | c_s[u]\eta_{is}(x) | \right\}^2 \leqq \left\{ \sum_{s=N_1}^{N} | \lambda_s | \frac{| \eta_{is}(x) |}{\lambda_s^2} | c_s[g] | \right\}^2$$
$$\leqq \left\{ \sum_{s=N_1}^{N} | \lambda_s | \left(\frac{\eta_{is}(x)}{\lambda_s^2} \right)^2 \right\} \left\{ \sum_{s=N_1}^{N} | \lambda_s | c_s^2[g] \right\}$$
$$\leqq l_{10} \left\{ \sum_{s=N_1}^{N} | \lambda_s | c_s^2[g] \right\}.$$

A corresponding inequality is obtained for the second series of (7.11). It therefore follows in view of Lemma 7.2 that these series converge absolutely and uniformly and the lemma is proven.

Since the series (7.7) converge absolutely and uniformly on x_1x_2 the functions $u_i(x) - \sum_{s=1}^{\infty} c_s[u]\eta_{is}(x)$ satisfy the relations (7.2), and therefore $K_{ij}\{u_j - \sum_{s=1}^{\infty} c_s[u]\eta_{js}\} \equiv 0$ on x_1x_2. In particular, if $| K_{ij} | \neq 0$ on x_1x_2 we have $u_i(x) = \sum_{s=1}^{\infty} c_s[u]\eta_{is}(x)$ $(i = 1, 2, \cdots, n)$.

From Theorem 7.1 and Lemma 7.3 we now obtain the following expansion theorem.

THEOREM 7.2. *Let g be an arbitrary admissible arc and u_i, ρ_a the solution of (7.3) corresponding to g. If v_i, σ_a is the solution of the system*

$$J_i(v, \sigma) + K_{ij}u_j = 0, \qquad \Phi_a(x, v, v') = 0,$$
$$L_{i1}(v, \sigma, d, 0) = 0 = L_{i2}(v, \sigma, d, 0) = \Psi_\gamma[v(x_1), v(x_2)],$$

then the series

$$\sum_{s=1}^{\infty} c_s[v]\eta_{is}(x), \qquad \sum_{s=1}^{\infty} c_s[v]\zeta_{is}(x)$$

converge absolutely and uniformly on the interval x_1x_2 to the values $v_i(x)$, $\Omega_{\eta'_i}[x, v(x), v'(x), \sigma(x)]$. The series $\sum_{s=1}^{\infty} c_s[v]\eta'_{is}(x)$ and $\sum_{s=1}^{\infty} c_s[v]\mu_{as}(x)$ also converge absolutely and uniformly to the functions $v'_i(x)$, $\sigma_a(x)$.

UNIVERSITY OF CHICAGO,
CHICAGO, ILL.

ON BOUNDARY VALUE PROBLEMS ASSOCIATED WITH DOUBLE INTEGRALS IN THE CALCULUS OF VARIATIONS.†

By WILLIAM T. REID.‡

1. *Introduction.* Lichtenstein § has considered a certain double integral problem with fixed boundary in the calculus of variations and has shown that the necessary condition of Jacobi is closely related to a certain boundary value problem associated with the Jacobi differential equation. If Jacobi's condition is satisfied, then the smallest positive characteristic number λ_1^+ of the boundary value problem which he considered must satisfy the inequality $\lambda_1^+ \geqq 1$, and conversely. To establish a certain minimizing property of λ_1^+ he has made use of some expansion theorems for arbitrary functions in terms of the characteristic solutions of the associated boundary value problem. It is the purpose of the present paper to establish by the use of the methods of the calculus of variations this minimizing property of the first characteristic number in a much more elementary manner than that given by Lichtenstein; in particular, the existence of λ_1^+ is established by the methods of the calculus of variations. It is shown that in the associated boundary value problem the parameter may be allowed to enter in a simpler form than that used by Lichtenstein. By the use of this minimizing property of the first characteristic number, sufficient conditions for an extremal surface to render the double integral a weak relative minimum may be readily given.

2. *A boundary value problem.* Let A be a simply connected region in xy-space which is bounded by a closed analytic curve C. Let P, Q, R, B and K be functions of (x, y) which are analytic in their arguments in a region of xy-space which contains $A + C$; it is also supposed that $PR - Q^2 > 0$, $P > 0$, and $B \leqq 0$ on $A + C$. In this section will be considered the boundary value problem

$$(1) \qquad \partial(P\zeta_x + Q\zeta_y)/\partial x + \partial(Q\zeta_x + R\zeta_y)/\partial y + (B + \lambda K)\zeta = 0,$$

$$(2) \qquad \zeta = 0 \text{ on } C.$$

† Presented to the American Mathematical Society, December 31, 1930.

‡ National Research Fellow in Mathematics.

§ L. Lichtenstein, *Monatshefte für Mathematik und Physik*, Vol. 28 (1917), pp. 3-51; also, *Mathematische Zeitschrift*, Vol. 6 (1919), pp. 26-51. These papers will be referred to as L. I and L. II respectively. See also, Picone, *Atti della Reale Accademia dei Lincei, Rendiconti* (5), Vol. 31_1 (1922), pp. 46-48.

A value λ will be said to be a characteristic number of this boundary value problem if corresponding to this value there exists a non-identically vanishing function $\zeta(x, y)$ which is of class C'' † on $A + C$, and satisfies the differential equation (1), together with the boundary condition (2).

For convenience, let

$$2\Omega(x, y, \zeta, \zeta_x, \zeta_y) = P\zeta_x^2 + 2Q\zeta_x\zeta_y + R\zeta_y^2 - B\zeta^2.$$

Then

$$2\Omega(x, y, \zeta, \zeta_x, \zeta_y) = \zeta_x\Omega_{\zeta_x} + \zeta_y\Omega_{\zeta_y} + \zeta\Omega_\zeta,$$
$$\eta_x\Omega_{\zeta_x} + \eta_y\Omega_{\zeta_y} + \eta\Omega_\zeta = \zeta_x\Omega_{\eta_x} + \zeta_y\Omega_{\eta_y} + \zeta\Omega_\eta, \ddagger$$

and the differential equations (1) may be written

$$(1') \qquad \partial\Omega_{\zeta_x}/\partial x + \partial\Omega_{\zeta_y}/\partial y - \Omega_\zeta + \lambda K\zeta = 0.$$

If now $I_2(\eta, \zeta)$ is defined as the double integral

$$(3) \qquad I_2(\eta, \zeta) = \int\!\!\int_A [\eta_x\Omega_{\zeta_x} + \eta_y\Omega_{\zeta_y} + \eta\Omega_\zeta]\, dx\, dy,$$

it follows after the usual integration by parts that

$$(4) \qquad I_2(\eta, \zeta) = \int_C \eta[\Omega_{\zeta_x}\, dy - \Omega_{\zeta_y}\, dx]$$
$$- \int\!\!\int_A \eta[\partial\Omega_{\zeta_x}/\partial x + \partial\Omega_{\zeta_y}/\partial y - \Omega_\zeta]\, dx\, dy.$$

If η and ζ are solutions of (1), (2) corresponding to distinct values λ and λ^*, since $I_2(\eta, \zeta) = I_2(\zeta, \eta)$, we have that

$$[\lambda - \lambda^*] \int\!\!\int_A K\eta\zeta dx\ dy = 0.$$

Since $I_2(\zeta) \equiv I_2(\zeta, \zeta) > 0$ for every function ζ which satisfies (2) and is not identically zero on A, it is then readily seen that all the characteristic

† A function of (x, y) will be said to be of class $C^{(n)}$ $(n = 1, 2, \cdots)$ on $A + C$ if the function and all its partial derivatives of order not greater than n are continuous in the region A and on its boundary C. A function is said to be of class D' on $A + C$ if it is continuous and furthermore A may be divided into a finite number of sub-regions each of which is bounded by a simple closed curve consisting of a finite number of analytic pieces, and in each sub-region the given function is identical with a function which is of class C' in that sub-region and on its boundary.

‡ The derivatives of Ω are understood to have the arguments $(\zeta, \zeta_x, \zeta_y)$ or (η, η_x, η_y) as indicated by their subscripts.

numbers of the boundary value problem (1), (2) are real. If ζ is a solution of (1), (2) corresponding to a value λ, relation (4) gives

$$I_2(\zeta) = \lambda \iint\limits_A K\zeta^2 \, dx \, dy, \tag{5}$$

and it is therefore seen that $\iint\limits_A K\zeta^2 \, dx \, dy$ has the sign of λ. If then $K(x, y) \leqq 0$ on $A + C$, the system (1), (2) can have no positive characteristic numbers. We will suppose in the following that $K(x, y) > 0$ for some point of A, and it will be shown that under this hypothesis there exists for (1), (2) a positive characteristic number.

Let H denote the class of functions $\zeta(x, y)$ which are of class D' on $A + C$ and vanish on C; H^* is then used to denote the class of functions ζ which belong to H and satisfy also the relation

$$\iint\limits_A K\zeta^2 \, dx \, dy = 1. \tag{6}$$

On the hypothesis that K is positive at some point of A, the class H^* is not empty. Now since $I_2(\zeta) > 0$ for all functions ζ of class H^*, there exists a non-negative greatest lower bound λ_1 of the values of $I_2(\zeta)$ in the class H^*. Then the expression

$$I_2(\zeta \mid \lambda_1) = I_2(\zeta) - \lambda_1 \iint\limits_A K\zeta^2 \, dx \, dy \tag{7}$$

is seen to be non-negative for all functions ζ of class H.

In this section will be established the following theorem:

THEOREM 2. 1. *If λ_1 is the greatest lower bound of $I_2(\zeta)$ in the class H^*, then $\lambda = \lambda_1$ is a characteristic number of the boundary value problem* (1), (2).

In the proof of this theorem the following lemma will be used.

LEMMA 2. 1. *There exists a positive constant α_1 such that if $\zeta(x, y)$ is any function of class D' on $A + C$ which vanishes on C, then*

$$\iint\limits_A [\zeta_x^2 + \zeta_y^2] \, dx \, dy \geqq \alpha_1 \iint\limits_A \zeta^2 \, dx \, dy. \tag{8}$$

For let S be a square which contains A in its interior and whose corners have the coördinates (x_0, y_0), $(x_0 + d, y_0)$, $(x_0 + d, y_0 + d)$, $(x_0, y_0 + d)$,

where $d > 0$. Let $\zeta(x, y)$ be defined as zero outside of A. Then by Schwarz' inequality,

$$[\zeta(x,y)]^2 = [\int_{x_0}^{x} \zeta_x(x,y)\,dx]^2 \leq [x - x_0] \int_{x_0}^{x} \zeta_x^2(x,y)\,dx$$
$$\leq [x - x_0] \int_{x_0}^{x_0+d} \zeta_x^2(x,y)\,dx$$

and therefore

$$\iint_A \zeta^2 dxdy = \iint_S \zeta^2 dxdy \leq (d^2/2) \iint_S \zeta_x^2 dxdy = (d^2/2) \iint_A \zeta_x^2 dxdy.$$

Inequality (8) then follows by combining this inequality and a similar one for $\iint_A \zeta_y^2\,dx\,dy$.

Theorem 2.1 will now be proved. The equation (1) for $\lambda = \lambda_1$ is the Euler-Lagrange equation for the problem of minimizing $I_2(\zeta \mid \lambda_1)$ in the class H. Since $I_2(\zeta \mid \lambda_1) \geq 0$, we have that $\zeta(x, y) \equiv 0$ is a minimizing surface for this integral, and therefore for it the necessary condition of Jacobi must be satisfied. That is, if $\bar{A}$ is a subregion of A bounded by a simply closed curve $\bar{C}$ which consists of a finite number of analytic pieces and lies in A, then there exists no solution of the Jacobi differential equation which vanishes on $\bar{C}$ and is not identically zero on $\bar{A}$.† Since the integrand of $I_2(\zeta \mid \lambda_1)$ is quadratic in ζ, ζ_x and ζ_y, we have that the Jacobi equation is identical with the Euler-Lagrange equation and is given by (1) with $\lambda = \lambda_1$.

Now if $\lambda = \lambda_1$ is not a characteristic number of (1), (2) there exists a unique solution $u(x, y)$ of (1) for $\lambda = \lambda_1$ such that $u(x, y) = 1$ on C.‡

† See Bolza, *Variationsrechnung*, Leipzig, 1909, p. 675. The proof of this condition depends upon the existence of an elementary solution of the Jacobi equation. The existence of an elementary solution for the general equation of elliptic type whose coefficients are analytic has been proven by Hadamard. See J. Hadamard, *Lectures on Cauchy's Problem*, London, 1923. The existence of an elementary solution for the general equation of elliptic type whose coefficients satisfy much weaker conditions has been established by Levi. See E. E. Levi, *Rendiconti del Circolo Matematico di Palermo*, Vol. 24 (1907), pp. 275-317.

‡ See L. I, p. 16. Lichtenstein shows that the boundary value problem (1), (2) may be reduced to one associated with the normal form of equation (1). The system (1), (2) may also be considered directly and it may be shown that if for $\lambda = \lambda_1$ there is no solution of the boundary value problem, then there exists a unique solution of (1) which coincides on the boundary C with an arbitrarily chosen analytic function. For a treatment of the corresponding problem for the general linear elliptic differential equation of the second order with three independent variables, see W. Sternberg,

It will now be shown that $u(x,y) > 0$ on A. For if at a point of A this solution $u(x,y)$ vanishes, then there are points of A at which $u(x,y) < 0$.† From this it follows that there is a sub-region $\bar{A}$ of A in which $u(x,y)$ is not identically zero and on the boundary $\bar{C}$ of $\bar{A}$ we have $u(x,y) = 0$. Furthermore, $\bar{C}$ consists of a finite number of analytic pieces.‡ But since λ_1 is the greatest lower bound of $I_2(\zeta)$ in the class H^*, this is impossible, as shown above. It therefore follows that $u(x,y) > 0$ on $A + C$.

If the transformation of Clebsch is then applied to $I_2(\zeta \mid \lambda_1)$, it is seen that for an arbitrary function $\zeta(x,y)$ of class H,§

$$(9) \qquad I_2(\zeta \mid \lambda_1) = \iint_A u^2 [P\bar{\zeta}_x^2 + 2Q\bar{\zeta}_x\bar{\zeta}_y + R\bar{\zeta}_y^2]\, dx\, dy,$$

where $\bar{\zeta} = \zeta/u$. Since $u(x,y) > 0$ on $A + C$, and $PR - Q^2 > 0$ on $A + C$, there is a positive constant α_2 such that

$$I_2(\zeta \mid \lambda_1) \geq \alpha_2 \iint_A [\bar{\zeta}_x^2 + \bar{\zeta}_y^2]\, dx\, dy,$$

and therefore, in view of lemma 2.1,

$$I_2(\zeta \mid \lambda_1) \geq \alpha_1\alpha_2 \iint_A \bar{\zeta}^2\, dx\, dy.$$

Now

$$\bar{\zeta}^2 \geq [\max(u)]^{-1}\zeta^2 \geq [\max(u)]^{-1}[\max(K)]^{-1}K\zeta^2,$$

and therefore there is a positive constant α_3 such that

$$I_2(\zeta \mid \lambda_1) \geq \alpha_3 \iint_A K\zeta^2\, dx\, dy.$$

But this inequality, which has been obtained on the assumption that λ_1 is not a characteristic number of (1), (2), is a contradiction to the hypothesis that λ_1 is the greatest lower bound of $I_2(\zeta)$ in the class H^*. Hence $\lambda = \lambda_1$

Mathematische Zeitschrift, Vol. 21 (1924), pp. 286-311. The coefficients of the equation treated by Sternberg are supposed to satisfy a condition much weaker than being analytic; in fact, if they are of class C''' then the condition imposed is satisfied.

† When the coefficients of (1) are analytic, as supposed above, this follows from results established by Picard. See Picard, *Traité d'Analyse*, Paris, 1905, Vol. 2, pp. 28-32. The result is also true when the coefficients of the differential equation are not required to be analytic; see L. Lichtenstein, *Rendiconti del Circolo Matematico di Palermo*, Vol. 33 (1912), pp. 201-211.

‡ See L. I, p. 10. The proof is there indicated for the normal form of (1), but the same proof may be applied directly to the differential equation (1).

§ Bolza, *loc. cit.*, p. 680.

is a characteristic number for the boundary value problem (1), (2). Furthermore, since (5) holds for every solution of (1) corresponding to a value λ, we have that $\lambda_1 > 0$ and that $\lambda = \lambda_1$ is the smallest positive characteristic number of (1), (2).

3. *The calculus of variations problem.* As in § 2, let A be a simply connected region in the xy-plane whose boundary C is a simple closed analytic curve. Consider the problem of minimizing the integral

$$I = \iint_A f(x, y, z, z_x, z_y)\,dx\,dy \tag{10}$$

in the class of surfaces

$$z = z(x, y) \tag{11}$$

which are of class D' on $A + C$ and which coincide with a given function $\psi(x, y)$ on the boundary C. The function $\psi(x, y)$ is supposed to be an analytic function of the arc element on C, and any surface z of the type described above will be said to be *admissible.*

Suppose that E is an admissible surface and in a neighborhood $\mathfrak{N}$ of the values (x, y, z, z_x, z_y) on E the function $f(x, y, z, p, q)$ is an analytic function of its five arguments. It is also supposed that E is an extremal surface, that is, it is of class C'' and satisfies the Euler-Lagrange differential equation

$$\partial f_p/\partial x + \partial f_q/\partial y - f_z = 0 \tag{12}$$

on $A + C$; furthermore, the Legendre condition is satisfied in the strong form, that is,

$$f_{pp}f_{qq} - f^2_{pq} > 0, \qquad f_{pp} > 0 \qquad [(x, y) \text{ on } A + C]. \tag{13}$$

In the partial derivatives of f which occur in (12) and (13) it is understood that the arguments $(x, y, z(x, y), z_x(x, y), z_y(x, y))$ are those which belong to E. According to Lichtenstein † the extremal surface $z = z(x, y)$ is then an analytic function of (x, y) on $A + C$.

The second variation of I along E is then given by

$$I_2(\zeta) = \iint_A 2\omega(x, y, \zeta, \zeta_x, \zeta_y)\,dx\,dy, \tag{14}$$

where $$2\omega(x, y, \zeta, \zeta_x, \zeta_y) = f_{pp}\zeta_x^2 + 2f_{pq}\zeta_x\zeta_y + f_{qq}\zeta_y^2 + 2f_{pz}\zeta_x\zeta + 2f_{qz}\zeta_y\zeta + f_{zz}\zeta^2.$$

† L. Lichtenstein, *Bulletin international de l'Académie des Sciences de Cracovie* (*Classe des sciences mathématiques et naturelles*) 1912, pp. 915-936.

A function $\zeta = \zeta(x, y)$ which is of class D' on $A + C$ and vanishes on C will be said to be an *admissible variation*. If $\zeta(x, y)$ is of class C'' on $A + C$, then by an integration by parts we obtain

$$I_2(\zeta) = -\iint_A \zeta J(\zeta)\, dx\, dy, \tag{15}$$

where

$$J(\zeta) = \partial\omega_{\zeta_x}/\partial x + \partial\omega_{\zeta_y}/\partial y - \omega. \tag{16}$$

It is seen that

$$J(\zeta) = L(\zeta) + k\zeta,$$

where

$$\begin{aligned} L(\zeta) &= \partial(f_{pp}\zeta_x + f_{pq}\zeta_y)/\partial x + \partial(f_{pq}\zeta_x + f_{qq}\zeta_y)/\partial y, \\ k(x, y) &= \partial f_{pz}/\partial x + \partial f_{qz}/\partial y - f_{zz}. \end{aligned} \tag{17}$$

If the terms of 2ω which contain $\zeta\zeta_x$ and $\zeta\zeta_y$ are integrated by parts, one obtains

$$I_2(\zeta) = \iint_A [f_{pp}\zeta_x^2 + 2f_{pq}\zeta_x\zeta_y + f_{qq}\zeta_y^2 - k\zeta^2]\, dx\, dy. \tag{14'}$$

Since $I_2(\zeta) \geqq 0$ for every admissible variation if E is a minimizing surface, we have: *If E is a minimizing surface for I, then there are no negative characteristic numbers for the boundary value problem*

$$J(\zeta) + \lambda\zeta = 0, \qquad \zeta = 0 \text{ on } C. \tag{18}$$

This follows immediately since if ζ is a solution of (18) corresponding to a value λ, then

$$I_2(\zeta) = \lambda \iint_A \zeta^2 dx\, dy.$$

The coefficients of the boundary value problem (18) as given above do not ssarily satisfy the conditions imposed on the coefficients of (1) in § 2, but positive number $\bar{\lambda}$ is chosen such that $|k(x, y)| \leqq \bar{\lambda}$ on $A + C$ and λ is ıced by $\mu = \lambda + \bar{\lambda}$, then (18) becomes

$$L(\zeta) + [(k - \bar{\lambda}) + \mu]\zeta = 0, \qquad \zeta = 0 \text{ on } C, \tag{}$$

the coefficients of (18′) satisfy the conditions imposed in § 2. The boun- value problem (18′) is seen to have only positive characteristic numbers nd by § 2 there exists a smallest characteristic number μ_1. Then $\mu_1 - \bar{\lambda}$ is the smallest characteristic number of (18) and we have seen if E is a minimizing surface, then $\lambda_1 \geqq 0$.

The following sufficiency condition will now be proved.

THEOREM 3.1. *If $z = z(x, y)$ is an analytic extremal surface along which the condition of Legendre is satisfied in the strong form and for which the smallest characteristic number of the boundary value problem* (18) *is positive, then z renders the integral I a weak relative minimum.*

For let $\zeta(x, y)$ be an arbitrary admissible variation such that the surface $z(x, y) + \zeta(x, y)$ lies in the neighborhood $\mathfrak{N}$ of the surface z. Then

$$\Delta I \equiv \iint_A f(x, y, z + \zeta, z_x + \zeta_x, z_y + \zeta_y)\, dx\, dy$$

$$- \iint_A f(x, y, z, z_x, z_y)\, dx\, dy$$

$$= \tfrac{1}{2} \iint_A 2\bar{\omega}(x, y, \zeta, \zeta_x, \zeta_y)\, dx\, dy,$$

where $2\bar{\omega}$ is a quadratic form in ζ, ζ_x, ζ_y whose coefficients are

$$\bar{f}_{pp} = 2 \int_0^1 (1 - \theta) f_{pp}(x, y, z + \theta\zeta, z_x + \theta\zeta_x, z_y + \theta\zeta_y)\, d\theta, \text{ etc.}$$

We will now say that an admissible variation ζ belongs to the class $R[\delta]$ if the functions $|\zeta|$, $|\zeta_x|$ and $|\zeta_y|$ are less than δ, uniformly on $A + C$. From the form of $2\bar{\omega}$ it is seen that for every positive ϵ there exists a positive δ_ϵ such that if ζ belongs to $R[\delta_\epsilon]$, then the coefficients of the quadratic form $2\bar{\omega} - 2\omega$ are all in absolute value less than ϵ, uniformly on $A + C$. Now apply the inequality $2ab \leq a^2 + b^2$ to each of the cross-product terms in the quadratic form $2\bar{\omega} - 2\omega$. By the use of Lemma 2.1, together with the fact that from (13) it follows that there is a positive constant α_4 such that

$$\zeta_x^2 + \zeta_y^2 \leq \alpha_4[f_{pp}\zeta_x^2 + 2f_{pq}\zeta_x\zeta_y + f_{qq}\zeta_y^2],$$

it is seen that if ζ belongs to $R[\delta_\epsilon]$, then

$$(22) \qquad \iint_A |\, 2\omega(x, y, \zeta, \zeta_x, \zeta_y) - 2\bar{\omega}(x, y, \zeta, \zeta_x, \zeta_y)\,|\, dx\, dy$$

$$\leq \epsilon\alpha_5 \iint_A [f_{pp}\zeta_x^2 + 2f_{pq}\zeta_x\zeta_y + f_{qq}\zeta_y^2]\, d$$

where $\alpha_5 = 3\alpha_4(1 + 1/\alpha_1)$. Let $\bar{\lambda}$ be a positive constant such that $|\, k($
$\leq \bar{\lambda}$ on $A + C$. It then follows from (14′) that if ζ belongs to $R[\delta_\epsilon]$,

$$\left| \iint_A [2\omega(x, y, \zeta, \zeta_x, \zeta_y) - 2\bar{\omega}(x, y, \zeta, \zeta_x, \zeta_y)]\,dx\,dy \right|$$
$$\leq \epsilon\alpha_5 \iint_A [2\omega(x, y, \zeta, \zeta_x, \zeta_y) + \bar{\lambda}\zeta^2]\,dx\,dy$$

and therefore

$$\Delta I \geq \tfrac{1}{2}\{(1 - \epsilon\alpha_5) \iint_A 2\omega(x, y, \zeta, \zeta_x, \zeta_y)\,dx\,dy - \epsilon\alpha_5\bar{\lambda} \iint_A \zeta^2 dx\,dy\}. \tag{23}$$

If now the smallest characteristic number λ_1 of (18) is positive, one may choose ϵ' such that

$$(1 - \epsilon'\alpha_5)\lambda_1 - \epsilon'\alpha_5\bar{\lambda} > 0,$$

and, furthermore, such that if ζ belongs to $R[\delta_{\epsilon'}]$, then the surface $z(x, y) + \zeta(x, y)$ lies in the neighborhood $\mathfrak{N}$ of the surface z. In view of the minimizing property of λ_1, it follows from (23) that if $\zeta(x, y)$ is an admissible variation which belongs to $R[\delta_{\epsilon'}]$, then

$$\Delta I \geq \tfrac{1}{2}[(1 - \epsilon'\alpha_5)\lambda_1 - \epsilon'\alpha_5\bar{\lambda}] \iint_A \zeta^2 dx\,dy, \tag{24}$$

and therefore for such variations ζ we have $\Delta I \geq 0$ and $\Delta I = 0$ only if $\zeta(x, y) \equiv 0$ on A. Theorem 3. 1 is therefore proved.

Instead of the boundary value problem (18) one may consider the boundary value problem

$$L(\zeta) + g\zeta + \lambda(k - g)\zeta = 0, \qquad \zeta = 0 \text{ on } C, \tag{25}$$

where $k(x, y)$ is defined by (17) and $g(x, y)$ is an arbitrary function which is analytic and non-positive on $A + C$; in terms of this boundary value problem one obtains then a sufficient condition for a weak relative minimum, analogous to Theorem 3. 1.

The boundary value problem (25) is the one considered by Lichtenstein. treated the particular case where $g(x, y) \equiv 0$, but he stated that results ogous to those established for this special case are true for the general lem (25). Since E is supposed to be an analytic extremal surface and function $f(x, y, z, p, q)$ is analytic in its arguments in a neighborhood $\mathfrak{N}$ he values corresponding to E, and since $g(x, y)$ is supposed to be analytic non-positive on $A + C$, the coefficients of the boundary value problem) satisfy the conditions imposed in § 2.

From the expression analogous to (5) we have that if ζ is a solution of) corresponding to a value λ, then λ and $\iint_A (k - g)\zeta^2 dx\,dy$ have the

same sign; furthermore, for such a solution ζ,

$$\iint_A 2\omega(x, y, \zeta, \zeta_x, \zeta_y)\,dx\,dy = (\lambda - 1)\iint_A (k - g)\zeta^2 dx\,dy.$$

From this relation, we have: *If E is a minimizing surface for I, then there exists no positive characteristic number of* (25) *which is less than unity.*

Corresponding to Theorem 3.1 we have the following sufficiency theorem.

THEOREM 3.2. *If $z = z(x, y)$ is an analytic extremal surface along which the condition of Legendre is satisfied in the strong form and for which the boundary value problem* (25) *has no positive characteristic number which is not greater than unity, then the surface z renders the integral I a weak relative minimum.*

For if $(k - g) > 0$ at a point of A, it then follows from § 2 that there is a first positive characteristic number λ_1^+ of (25), and for all admissible variations ζ we have the relation

$$(26)\qquad \iint_A [f_{pp}\zeta_x^2 + 2f_{pq}\zeta_x\zeta_y + f_{qq}\zeta_y^2 - g\zeta^2]\,dx\,dy \\ - \lambda_1^+ \iint_A (k - g)\zeta^2\,dx\,dy \geq 0.$$

From (14′) and (26) it then follows that

$$(27)\qquad \iint_A 2\omega(x, y, \zeta, \zeta_x, \zeta_y)\,dx\,dy \\ \geq (1 - 1/\lambda_1^+)\iint_A [f_{pp}\zeta_x^2 + 2f_{pq}\zeta_x\zeta_y + f_{qq}\zeta_y^2 - g\zeta^2]\,dx\,dy.$$

Since $g(x, y) \leq 0$, it follows from (22) that if ζ belongs to $R[\delta_\epsilon]$, then

$$(28)\qquad \Delta I \geq \tfrac{1}{2}(1 - 1/\lambda_1^+ - \epsilon\alpha_5) \\ \iint_A [f_{pp}\zeta_x^2 + 2f_{pq}\zeta_x\zeta_y + f_{qq}\zeta_y^2 - g\zeta^2]\,dx\,dy.$$

If $\lambda_1^+ > 1$, then ϵ' may be chosen so small that $1 - 1/\lambda_1^+ - \epsilon'\alpha_5 > 0$ and also such that if $\zeta(x, y)$ belongs to $R[\delta_{\epsilon'}]$ then the surface $z(x, y) + \zeta(x, y)$ lies in the neighborhood $\mathfrak{N}$ of the surface $z(x, y)$. It then follows that if $\zeta(x, y)$ is an admissible variation which belongs to $R[\delta_{\epsilon'}]$ then $\Delta I \geq 0$ and the equality sign holds only if $\zeta(x, y) \equiv 0$ on A. Theorem 3.2 is therefore proved in case $(k - g) > 0$ at some point of A. If $(k - g) \leq 0$ on A, then